

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2003, 09:34:45 ; Search time 110 Seconds
(without alignments)
2588.701 Million cell updates/sec

Title: US-10-087-782A-31
Perfect score: 7071
Sequence: 1 MTRKRTYWVPNSSGGLVNRG.....RKKPGSLEAALMATATSSLR 1382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				SUMMARIES			
Result No.	Score	Query Match	Length	DB ID	Description		
1	7071	100.0	1382	4	Q96J66	Q96j66 homo sapien	
2	7063	99.9	1382	4	Q96JA6	Q96ja6 homo sapien	
3	7028	99.4	1382	4	Q9BX80	Q9bx80 homo sapien	
4	6841	96.7	1344	4	Q8TDJ0	Q8tdj0 homo sapien	
5	3331	47.1	1359	4	Q96J65	Q96j65 homo sapien	
6	2251.5	31.8	1009	4	Q8TAF0	Q8taf0 homo sapien	
7	1992	28.2	1390	5	Q8T6H6	Q8t6h6 dictyosteli	
8	1980.5	28.0	935	4	Q8TEY2	Q8tey2 homo sapien	
9	1946.5	27.5	1334	5	Q8ST87	Q8st87 dictyosteli	
10	1946	27.5	1427	5	O62170	O62170 caenorhabdi	
11	1918	27.1	1264	5	Q8T6G8	Q8t6g8 dictyosteli	
12	1908	27.0	1323	5	Q8T1F1	Q8t1f1 dictyosteli	
13	1907.5	27.0	1273	13	Q8QFL6	Q8qfl6 brachydanio	
14	1903.5	26.9	1247	5	Q8T6H5	Q8t6h5 dictyosteli	
15	1848.5	26.1	1593	5	Q8T6H1	Q8t6h1 dictyosteli	
16	1816.5	25.7	1544	6	Q95M36	Q95m36 canis famil	

17	1805	25.5	1548	5	Q8T9C5	Q8t9c5 drosophila
18	1797.5	25.4	1564	13	Q8QG98	Q8qg98 raja erinac
19	1797.5	25.4	1622	10	Q9C8G9	Q9c8g9 arabidopsis
20	1797.5	25.4	1623	10	O48907	O48907 arabidopsis
21	1795.5	25.4	1623	10	O64590	O64590 arabidopsis
22	1795	25.4	1495	10	Q9C8H0	Q9c8h0 arabidopsis
23	1794.5	25.4	1528	11	O35379	O35379 mus musculu
24	1793.5	25.4	1622	10	O24635	O24635 arabidopsis
25	1784	25.2	1514	4	Q96QA9	Q96qa9 homo sapien
26	1781.5	25.2	1534	5	Q95QE2	Q95qe2 caenorhabdi
27	1778.5	25.2	1623	10	O22449	O22449 arabidopsis
28	1778	25.1	1622	10	O48908	O48908 arabidopsis
29	1777.5	25.1	1528	5	Q9N2N3	Q9n2n3 caenorhabdi
30	1771.5	25.1	1543	11	O8VI47	O8vi47 mus musculu
31	1770.5	25.0	1540	5	Q94136	Q94136 caenorhabdi
32	1770.5	25.0	1543	11	Q8VI46	Q8vi46 mus musculu
33	1767	25.0	1515	4	Q9UQ99	Q9uq99 homo sapien
34	1766.5	25.0	1547	5	Q9I7N0	Q9i7n0 drosophila
35	1766.5	25.0	1591	5	Q9VK56	Q9vk56 drosophila
36	1763.5	24.9	1573	5	Q93552	Q93552 caenorhabdi
37	1762	24.9	1525	5	Q20943	Q20943 caenorhabdi
38	1760	24.9	1493	10	Q94E55	Q94e55 oryza sativ
39	1760	24.9	1544	6	Q95L75	Q95l75 macaca mula
40	1756.5	24.8	1502	6	Q9GK09	Q9gk09 canis famil
41	1755.5	24.8	1490	10	Q9M1C7	Q9mlc7 arabidopsis
42	1755.5	24.8	2580	5	Q9VJ21	Q9vj21 drosophila
43	1754.5	24.8	1543	11	Q9JIH8	Q9jih8 mus musculu
44	1754	24.8	1525	5	Q94137	Q94137 caenorhabdi
45	1743.5	24.7	1368	10	Q9S9R0	Q9s9r0 arabidopsis

ALIGNMENTS

RESULT 1
Q96J66
ID Q96J66 PRELIMINARY; PRT; 1382 AA.
AC Q96J66;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ATP-binding cassette transporter sub-family C member 11.
GN ABCC11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=21376129; PubMed=11483364;
RA Tammur J., Prades C., Arnould I., Rzhetsky A., Hutchinson A.,
RA Adachi M., Schuetz J.D., Swoboda K.J., Ptacek L.J., Rosier M.,
RA Dean M., Allikmets R.;
RT "Two new genes from the human ATP-binding cassette transporter
superfamily, ABCC11 and ABCC12, tandemly duplicated on chromosome
16q12.";
RL Gene 273:89-96(2001).
DR EMBL; AY040219; AAK76739.1; -.
DR InterPro; IPR001140; ABCtranprtrTM.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transportr; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_2.
KW ATP-binding.
SQ SEQUENCE 1382 AA; 154300 MW; 0F3C17DC69AE97F4 CRC64;

Query Match 100.0%; Score 7071; DB 4; Length 1382;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTRKRTYWVPNSSGGLVNRGIDGDDMVSGGLIYKTYTLQDGPWSQQERNPEAPGRAAVPP 60

Db 1 MTRKRTYVVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQQERNPEAPGRAAVPP 60
QY 61 WGKYDAALRTMIPFRPKPRFPAPQPLDNAGLFSYLTVSWLTPLMIQSLRSLDENTIPPL 120
Db 61 WGKYDAALRTMIPFRPKPRFPAPQPLDNAGLFSYLTVSWLTPLMIQSLRSLDENTIPPL 120
QY 121 SVHDASDKNVQRLHRLWEEEVSRRGIEKASVLLVMLRFQTRRLIFDALLGICFCIASVLG 180
Db 121 SVHDASDKNVQRLHRLWEEEVSRRGIEKASVLLVMLRFQTRRLIFDALLGICFCIASVLG 180
QY 181 PILIIPKILEYSEEQLGNVHVHGLCFALFLSECVKSLSFSSSWIINQRTAIRFRAAVSS 240
Db 181 PILIIPKILEYSEEQLGNVHVHGLCFALFLSECVKSLSFSSSWIINQRTAIRFRAAVSS 240
QY 241 FAFEKLIQFKSVIHITSGEAISFTTGDVNYLFEVCYGPLVLITCASLVICSISSYFIIG 300
Db 241 FAFEKLIQFKSVIHITSGEAISFTTGDVNYLFEVCYGPLVLITCASLVICSISSYFIIG 300
QY 301 YTAFAIILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQIRVITSEVLTICIKLIKMYTWEKP 360
Db 301 YTAFAIILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQIRVITSEVLTICIKLIKMYTWEKP 360
QY 361 FAKIILEDLRRKERKLLKCKGLVQSLTSITLFIPTVATAVWVLIHTSLKCLKLTASMAFSM 420
Db 361 FAKIILEDLRRKERKLLKCKGLVQSLTSITLFIPTVATAVWVLIHTSLKCLKLTASMAFSM 420
QY 421 LASLNLRLSVFEFVPIAVKGLTNSKSAVMRFKFFLQESPVFYVQTLQDPSKALVFEET 480
Db 421 LASLNLRLSVFEFVPIAVKGLTNSKSAVMRFKFFLQESPVFYVQTLQDPSKALVFEET 480
QY 481 LSWQQTCPGIVNGALELERNGHASEGMPTRPRDALGPREEGNSLGPDLHKINLVVSKGMML 540
Db 481 LSWQQTCPGIVNGALELERNGHASEGMPTRPRDALGPREEGNSLGPDLHKINLVVSKGMML 540
QY 541 GVCNGTSGKSSLLSAILEEMHLLGSGVQGSLAYVPPQAWIVSGNIRENILMGGAYDK 600
Db 541 GVCNGTSGKSSLLSAILEEMHLLGSGVQGSLAYVPPQAWIVSGNIRENILMGGAYDK 600
QY 601 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKORISLARAVYSRQIYLLDDP 660
Db 601 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKORISLARAVYSRQIYLLDDP 660
QY 661 LSAVDAHVGKHIFEICIKKTLRGKTVVLVTHQLQYLEFCGQIILLKNGKICENGTHSELM 720
Db 661 LSAVDAHVGKHIFEICIKKTLRGKTVVLVTHQLQYLEFCGQIILLKNGKICENGTHSELM 720
QY 721 QKKGKYAQLIQKMHKEATSDMLQDTAKIAEKPKVESQALATSLEESLNGNAVPEHQLTQE 780
Db 721 QKKGKYAQLIQKMHKEATSDMLQDTAKIAEKPKVESQALATSLEESLNGNAVPEHQLTQE 780
QY 781 EEMEEGSLSWRVYHHYIQAAGGYMVSCIIFFVVLIVFLTIFSFVWLSYWLEQSGGTNSS 840
Db 781 EEMEEGSLSWRVYHHYIQAAGGYMVSCIIFFVVLIVFLTIFSFVWLSYWLEQSGGTNSS 840
QY 841 RESNGTMADLGNADNPQLSFYQLVYGLNALLLICVGVCSGGIFTKVTRKASTALHNKLF 900
Db 841 RESNGTMADLGNADNPQLSFYQLVYGLNALLLICVGVCSGGIFTKVTRKASTALHNKLF 900
QY 901 NKVFRCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLLIVSVLSP 960
Db 901 NKVFRCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLLIVSVLSP 960
QY 961 YILLMGAILMVICFIYMMFKKAIGVFKRLENYSRSPFLSHILNSLOGLSSIHVYGKTED 1020
Db 961 YILLMGAILMVICFIYMMFKKAIGVFKRLENYSRSPFLSHILNSLOGLSSIHVYGKTED 1020
QY 1021 FISQFKRLTDAQNNYLLLLFLSSSTRWMAIRLEIMTNLVTLAVALFAFGISSTPYSEKUMA 1080
Db 1021 FISQFKRLTDAQNNYLLLLFLSSSTRWMAIRLEIMTNLVTLAVALFAFGISSTPYSEKUMA 1080
QY 1081 VNIVLQLASSFOATARIGLETEAQFTAVERILQYMKMCVSEAPLHMEGTSCPQGWPHGE 1140

Db 1081 VNIVLQLASSFOATARIGLETEAQFTAVERILQYMKMCVSEAPLHMEGTSCPQGWPHGE 1140
QY 1141 IIFQDYHMKYRDNTPTVLHGINTLIRGHEVVGIVGRTSGKSSGLGMAFLRFLVEPMAGRIL 1200
Db 1141 IIFQDYHMKYRDNTPTVLHGINTLIRGHEVVGIVGRTSGKSSGLGMAFLRFLVEPMAGRIL 1200
QY 1201 IDGVDICSIGLEDLRSKLSVIPQDPVLLSCTIRFNLDPFDRHTDQOIWDALERTFLTAKI 1260
Db 1201 IDGVDICSIGLEDLRSKLSVIPQDPVLLSCTIRFNLDPFDRHTDQOIWDALERTFLTAKI 1260
QY 1261 SKFPKKLHTDVVENGNGFVSGERQLLCITARAVLRNSKIILIDEATASIDMETDTLIQRTI 1320
Db 1261 SKFPKKLHTDVVENGNGFVSGERQLLCITARAVLRNSKIILIDEATASIDMETDTLIQRTI 1320
QY 1321 REAFQGCCTVLVIAHRVTTVLNCDHILVMNGKVVEFDRPEVLRKKPGSLFAALMATATSS 1380
Db 1321 REAFQGCCTVLVIAHRVTTVLNCDHILVMNGKVVEFDRPEVLRKKPGSLFAALMATATSS 1380
QY 1381 LR 1382
Db 1381 LR 1382
RESULT 2
Q96JA6 PRELIMINARY; PRT; 1382 AA.
AC Q96JA6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ATP-binding cassette protein C11.
GN ABCC11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Yabuuchi H., Ishikawa T.;
RT "Molecular cloning and characterization of human ABCC11 cDNA."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF367202; AAK58869.1; -
DR InterPro; IPR001140; ABCtranprtrTM.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transportr; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_2.
KW ATP-binding.
SQ SEQUENCE 1382 AA; 154273 MW; B4C220502A7001E5 CRC64;
Query Match 99.9%; Score 7063; DB 4; Length 1382;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1380; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MTRKRTYVVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQQERNPEAPGRAAVPP 60
Db 1 MTRKRTYVVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQQERNPEAPGRAAVPP 60
QY 61 WGKYDAALRTMIPFRPKPRFPAPQPLDNAGLFSYLTVSWLTPLMIQSLRSLDENTIPPL 120
Db 61 WGKYDAALRTMIPFRPKPRFPAPQPLDNAGLFSYLTVSWLTPLMIQSLRSLDENTIPPL 120
QY 121 SVHDASDKNVQRLHRLWEEEVSRRGIEKASVLLVMLRFQTRRLIFDALLGICFCIASVLG 180
Db 121 SVHDASDKNVQRLHRLWEEEVSRRGIEKASVLLVMLRFQTRRLIFDALLGICFCIASVLG 180
QY 181 PILIIPKILEYSEEQLGNVHVHGLCFALFLSECVKSLSFSSSWIINQRTAIRFRAAVSS 240
Db 181 PILIIPKILEYSEEQLGNVHVHGLCFALFLSECVKSLSFSSSWIINQRTAIRFRAAVSS 240
QY 241 FAFEKLIQFKSVIHITSGEAISFTTGDVNYLFEVCYGPLVLITCASLVICSISSYFIIG 300

QY	421	LASLNLLRLSVFFVPIAVKGLTNSKSAVMRFKKFFLQESPVFYVQTLQDPSKALVFEEAT	480
Db			
QY	421	LASLNLLRLSVFFVPIAVKGLTNSKSAVMRFKKFFLQESPVFYVQTLQDPSKALVFEEAT	480
Db			
QY	481	LSWQQTCPGIVNGALELERNGHASEGMRPRDALGP EEGNSLGP ELHKINLVVSKGMML	540
Db			
QY	481	LSWQQTCPGIVNGALELERNGHASEGMRPRDALGP EEGNSLGP ELHKINLVVSKGMML	540
Db			
QY	541	GVCNGTSGKSSLLSAILEEMHLLGSGVQGSLAYVPQQAIVSGNIRENILMGAYDK	600
Db			
QY	541	GVCNGTSGKSSLLSAILEEMHLLGSGVQGSLAYVPQQAIVSGNIRENILMGAYDK	600
QY	601	ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKQRI SARAVYSDRQIYLLDDP	660
Db			
QY	601	ARYLQVLHCCSLNRDLELLPFGDMTEIGERGNLGGQKQRI SARAVYSDRQIYLLDDP	660
QY	661	LSAVDAHVGKHIFEECIKKTLRGKTVVLVTHQLQYLEFCGQIILLKENGKICENGTHSELM	720
Db			
QY	661	LSAVDAHVGKHIFEECIKKTLRGKTVVQVTHQLQYLEFCGQVILLKENGKICENGTHSELM	720
QY	721	QKKGKYAQLIQMHKEATSDMLQDTAKIAEKP KVESQALATSLEESLNGNAVPEHQLTQE	780
Db			
QY	721	QKKGKYAQLIQMHKEATSDMLQDTAKIAEKP KVESQALATSLEESLNGNAVPEHQLTQE	780
QY	781	EEMEEGSLSWRVYHHYIQ AAGGYMVSCIIFFFWLVIVFLTIFSFWMLSYWLEQSGGTNSS	840
Db			
QY	781	EEMEEGSLSWRVYHHYIQ AAGGYMVSCIIFFFWLVIVFLTIFSFWMLSYWLEQSGGTNSS	840
QY	841	RESNGTMADLGNADNPQLSFYQLVYGLNALLICVGCSSGIFTKVRKASTALHNKLF	900
Db			
QY	841	RESNGTMADLGNADNPQLSFYQLVYGLNALLICVGCSSGIFTKVRKASTALHNKLF	900
QY	901	NKVERCPMSFDTPIGRLLNCFAGDLEQLDQLLP IFSEQFLVLSLMI AVLLIVSVLSP	960
Db			
QY	901	NKVERCPMSFDTPIGRLLNCFAGDLEQLDQLLP IFSEQFLVLSLMI AVLLIVSVLSP	960
QY	961	YILLMGAIIMVICFIYYMMFKKAIGVFKRLENYSRSP LFSHILNSLQGLSSIHVYGKTED	1020
Db			
QY	961	YILLMGAIIMVICFIYYMMFKKAIGVFKRLENYSRSP LFSHILNSLQGLSSIHVYGKTED	1020
QY	1021	FISQFKRLTDAQNNYLLLLFLSSSTRWMA LRLEIMTNLVT LAVALFVAFGISSTPY SFKVMA	1080
Db			
QY	1021	FISQFKRLTDAQNNYLLLLFLSSSTRWMA LRLEIMTNLVT LAVALFVAFGISSTPY SFKVMA	1080
QY	1081	VNIVLQASSFQATARIGLETEAQFT AVERILO YMKMCVSEAPLHMEGTS CPQGWPHGE	1140
Db			
QY	1081	VNIVLQASSFQATARIGLETEAQFT AVERILO YMKMCVSEAPLHMEGTS CPQGWPHGE	1140
QY	1141	IIFQDYHMKYRDNTPTVLHG INLTIRGHEVVGIVGRTGSGKSSLGMA LFR LVEPMA GRIL	1200
Db			
QY	1141	IIFQDYHMKYRDNTPTVLHG INLTIRGHEVVGIVGRTGSGKSSLGMA LFR LVEPMA GRIL	1200
QY	1201	IDGVDICSIGLEDLRSKLSVIPQDPVLLSGTIRFNLD PFD RHTDQOIWDALERTFLTKAI	1260
Db			
QY	1201	IDGVDICSIGLEDLRSKLSVIPQDPVLLSGTIRFNLD PFD RHTDQOIWDALERTFLTKAI	1260
QY	1261	SKFPKKLHTDVVENGGNFSVGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTI	1320
Db			
QY	1261	SKFPKKLHTDVVENGGNFSVGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTI	1320
QY	1321	REAFQGCCTVLVIAHRVTTVLNCDHILVMGNGKVVEFDRPEVLRKKPGSLFAALMATATSS	1380
Db			
QY	1321	REAFQGCCTVLVIAHRVTTVLNCDRILVMGNGKVVEFDRPEVLRKKPGSLFAALMATATSS	1380
QY	1381	LR 1382	
Db			
QY	1381	LR 1382	

RESULT 4
Q8TDJ0
ID Q8TDJ0
AC Q8TDJ0;

DT	01-JUN-2002	(TrEMBLrel. 21, Created)
DT	01-JUN-2002	(TrEMBLrel. 21, Last sequence update)
DT	01-JUN-2002	(TrEMBLrel. 21, Last annotation update)
DE	ATP-binding cassette protein C11 isoform A.	
GN	ABCC11.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=LIVER;	
RX	MEDLINE=21547789; PubMed=11688999;	
RA	yabuuchi H., Shimizu H., Takayanagi S., Ishikawa T.;	
RT	"Multiple splicing variants of two new human ATP-binding cassette	
RT	transporters, ABCC11 and ABCC12.";	
RL	Biochem. Biophys. Res. Commun. 288:933-939(2001).	
DR	EMBL; AF411579; AAL99902.1; -.	
KW	ATP-binding.	
SQ	SEQUENCE 1344 AA; 150065 MW; CEFFDBA8DADA4DAB CRC64;	
Query Match		
Best Local Similarity 96.7%; Score 6841; DB 4; Length 1344;		
Matches 1342; Conservative 1; Mismatches 1; Indels 38; Gaps 1;		
QY	1	MTRKRTYWVPNSSGGLVNRGIDIGDDMVSGLIYKTYTTLQDGPWSQQERNPEAPRAAVPP 60
Db		
QY	61	WGKYDAALRTMIPFRPKRPAPQPLDNAGLFSYLT VSWLTPLMIQSLRSRLDENTIPPL 120
Db		
QY	61	WGKYDAALRTMIPFRPKRPAPQPLDNAGLFSYLT VSWLTPLMIQSLRSRLDENTIPPL 120
QY	121	SVHDASDNVQRLHRLWEEVSRRGIEKASVLLVLMRLRFQRTLIFDALLGICFCIASVLG 180
Db		
QY	121	SVHDASDNVQRLHRLWEEVSRRGIEKASVLLVLMRLRFQRTLIFDALLGICFCIASVLG 180
QY	181	PILIIPKILEYSEELGNVHVHVGCLPALFLSECVKSLSFSSSWIINQRTAIRFRAAVSS 240
Db		
QY	181	PILIIPKILEYSEELGNVHVHVGCLPALFLSECVKSLSFSSSWIINQRTAIRFRAAVSS 240
QY	241	FAFEKLIQPKSVIHITSGEAI SFFTG DVNVLFE GVCYGPLVLITCASLVICSISYFIIG 300
Db		
QY	241	FAFEKLIQPKSVIHITSGEAI SFFTG DVNVLFE GVCYGPLVLITCASLVICSISYFIIG 300
QY	301	YTAFAIALCYLLVFPLAVFMT RMAVKAQHTSEVSDQRI RVTSEVLT CIKLIKMYTWEKP 360
Db		
QY	301	YTAFAIALCYLLVFPLAVFMT RMAVKAQHTSEVSDQRI RVTSEVLT CIKLIKMYTWEKP 360
QY	361	FAKIIEDLRKERRKLLLEKCGLVQSLTSITLFI IPTVATAVWVLIHTSLK LKLTASMAFSM 420
Db		
QY	361	FAKIIEDLRKERRKLLLEKCGLVQSLTSITLFI IPTVATAVWVLIHTSLK LKLTASMAFSM 420
QY	421	LASNLRLSVFFVPIAVKGLTNSKSAVMRFKKFFLQESPVFYVQTLQDPSKALVFEEAT 480
Db		
QY	421	LASNLRLSVFFVPIAVKGLTNSKSAVMRFKKFFLQESPVFYVQTLQDPSKALVFEEAT 480
QY	481	LSWQQTCPGIVNGALELERNGHASEGMRPRDALGP EEGNSLGP ELHKINLVVSKGMML 540
Db		
QY	481	LSWQQTCPGIVNGALELERNGHASEGMRPRDALGP EEGNSLGP ELHKINLVVSKGMML 540
QY	541	GVCNGTSGKSSLLSAILEEMHLLGSGVQGSLAYVPQQAIVSGNIRENILMGAYDK 600
Db		
QY	541	GVCNGTSGKSSLLSAILEEMHLLGSGVQGSLAYVPQQAIVSGNIRENILMGAYDK 600
QY	601	ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKQRI SARAVYSDRQIYLLDDP 660
Db		
QY	601	ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKQRI SARAVYSDRQIYLLDDP 660
QY	661	LSAVDAHVGKHIFEECIKKTLRGKTVVLVTHQLQYLEFCGQIILLKENGKICENGTHSELM 720
Db		
QY	661	LSAVDAHVGKHIFEECIKKTLRGKTVVLVTHQLQYLEFCGQIILLKENGKICENGTHSELM 720

QY	721	QKKGKYAQLIQMHKEATSDMLQDTAKIAEKPKVESQALATSLAESLNGNAVPEHQLTQE	780
Db	721	QKKGKYAQLIQMHKEATSDMLQDTAKIAEKPKVESQALATSLAESLNGNAVPEHQLTQE	780
QY	781	EEMEEGSLSWRVYHHYIOAAGGYMVSCIIFFVVLIVFLTFPSFWLSYWLQGGSGTNSS	840
Db	781	EEMEEGSLSWRVYHHYIOAAGGYMVSCIIFFVVLIVFLTFPSFWLSYWLQGGSGTNSS	840
QY	841	RESNGTMADLGNADNPQLSFYQLVYGLNALLICVGCSSGIFTKVTRKASTALHNKLF	900
Db	841	RESNGTMADLGNADNPQLSFYQLVYGLNALLICVGCSSGIFTKVTRKASTALHNKLF	900
QY	901	NKVERCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLWIAVLLIVSVLSP	960
Db	901	NKVERCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLWIAVLLIVSVLSP	960
QY	961	YILLMGATIMVICFIYYMFKKATGVFKRLNYSRSPLEFSLNLSLQGLSSIHVYKTED	1020
Db	961	YILLMGATIMVICFIYYMFKKATGVFKRLNYSRSPLEFSLNLSLQGLSSIHVYKTED	1020
QY	1021	FISQFKRLTDAQNYYLLFLSSTRWMLRLEIMTNLVTLAVALFVAFGISSTPYSEKMA	1080
Db	1021	FISQFKRLTDAQNYYLLFLSSTRWMLRLEIMTNLVTLAVALFVAFGISSTPYSEKMA	1080
QY	1081	VNIVLQLASSFQATARIGLETEAQFTAVERILQYMKMCVSEAPLHMEGTSCPGWPQHGE	1140
Db	1081	VNIVLQLASSFQATARIGLETEAQFTAVERILQYMKMCVSEAPLHMEGTSCPGWPQHGE	1140
QY	1141	IIFQDYHMKYRDNTPTVLHGINLTIRGHEVVGIVGRTGSGKSSGLMALFRLVEPMAGRIL	1200
Db	1141	IIFQDYHMKYRDNTPTVLHGINLTIRGHEVVGIVGRTGSGKSSGLMALFRLVEPMAGRIL	1200
QY	1201	IDGVDICSIGLEDRLSKLSVIPQDPVLLSGTIRFNLDPFDRHTDQIWDALERTFLPKAI	1260
Db	1201	IDGVDICSIGLEDRLSKLSVIPQDPVLLSGTIRFNLDPFDRHTDQIWDALERTFLPKA	1259
QY	1261	SKFPKKLHTDVVENGNGFSVGERQLLCIARAVLRNSKIILLIDEATASIDMETDTLIQRTI	1320
Db	1260	-----IILLIDEATASIDMETDTLIQRTI	1282
QY	1321	REAFQGCTVLVIAHRVTTVLNCDHILVMNGKVVFEFDRPEVLRRKKPGSLFAALMATATSS	1380
Db	1283	REAFQGCTVLVIAHRVTTVLNCDHILVMNGKVVFEFDRPEVLRRKKPGSLFAALMATATSS	1342
QY	1381	LR 1382	
Db	1343	LR 1344	

RESULT 5

Q96J65			
ID	Q96J65	PRELIMINARY;	PRT; 1359 AA.
AC	Q96J65;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	ATP-binding cassette transporter sub-family C member 12.		
GN	ABCC12.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=TESTIS;		
RX	MEDLINE=21376129; PubMed=11483364;		
RA	Tammur J., Prades C., Arnould I., Rzhetsky A., Hutchinson A.,		
RA	Adachi M., Schuetz J.D., Swoboda K.J., Ptacek L.J., Rosier M.,		
RA	Dean M., Allikmets R.;		
RT	"Two new genes from the human ATP-binding cassette transporter		
RT	superfamily, ABCC11 and ABCC12, tandemly duplicated on chromosome		
RT	16q12.";		
RL	Gene 273:89-96(2001).		

DR	EMBL; AY040220; AAK76740.1; -.		
DR	InterPro; IPR001140; ABCtranprtrTM.		
DR	InterPro; IPR003439; ABC_transportr.		
DR	Pfam; PF00664; ABC_membrane; 2.		
DR	Pfam; PF00005; ABC_tran; 2.		
DR	ProDom; PD000006; ABC_transportr; 2.		
DR	PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_2.		
KW	ATP-binding; GTP-binding; Protein biosynthesis.		
SQ	SEQUENCE 1359 AA; 152243 MW; 2A8AF2A535F6AA85 CRC64;		
	Query Match	47.1%;	Score 3331; DB 4; Length 1359;
	Best Local Similarity	48.6%;	Pred. No. 2e-231;
	Matches	653; Conservative	243; Mismatches 404; Indels 44; Gaps 7;
QY	63	KYDAALRTMIPFRPKPRFPAPQPLDNAGLFSYLTVSWLTPLMIQSLRSRLDENTIPPLSV	122
		: : : : : : : : : :	
Db	24	RYDPSLKTMIIPVPCARL-APNPVDDAGLLSFATFSWLTPEVMVKGYRQRLTVDTLPPLST	82
QY	123	HDASDKNVQRLHRLWEEVESVRGIEKASVLLVMLRFQRTLRIFDALLGICFCIASVLGPI	182
		: : : : : : : : :	
Db	83	YDSDTNAKRFRVLWDEVARVGPKEKASLSHVVKFQRTRLVMDIVANILCIIMAAIGPV	142
QY	183	LIIPKILEYSEEQLGNVVHVGVLGFALFLSECVKSLSFSSSWIINORTAIRFAAVSSFA	242
		: : : : : : : : :	
Db	143	ILIHQILQQTERTSGKVWVGIGLCIALFATEFTKVFFWALAWAINYRTAIRLKVALSTLV	202
QY	243	FEKLITQFKSVIHITSGEAISFTGDVNYLFEVCYGPLVLITCASLVICSISSYFIIGYT	302
		: : : : : : : : : : : : : :	
Db	203	FENLVSFKTLTHISVGEVLNLLSSDSYSLFEAALFCPLPATIPLMVFCAAYAFFILGPT	262
QY	303	AFTAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQRIRVTSVLTCTIKLIKMYTWKPPFA	362
		: : : : : : : : : : : : : :	
Db	263	ALIGTSVYVIFIPVQMFMAKLSAFRRSAILVTDKRVQTMNEFLTCTIRLIKMYAWEKSFT	322
QY	363	KIITEDLRKKERKLLKCKGLVQSITSITLFIPTVATAVWVLHTSLKCLKLTASMAFSMLA	422
		: : : : : : : : : : :	
Db	323	NTIQDIRRRERKLLKAGFVQSGNSALAPIVSTIAIVLTLSCHILLRRKLTAPVAFSVIA	382
QY	423	SLNLLRLSVFFVPIAVKGLTNSKSAVMRFKKFFLOESPVFVQTLQDPSKALVFEEATLS	482
		: : : : : : : : : : : : : :	
Db	383	MFNVMKFSIAILPFSIKAMAEANVSLRRMKKILIDKSPPSYITQPEDPDTVLLIANATLT	442
QY	483	WQ-----QTCPGIVNGALEL---ERNGHASEGMRPRDRLGPEEGNSLGPDLHKINLV	533
		: : : : : : : : : : : : : : : : : :	
Db	443	WEHEASRKSTPKKLQNKRHLCCKQRSEAYSERSPPAKGATGPEEQSDSLKSVLHSISFV	502
QY	534	VSKGMMLGVCNGTSGSKSSLISAILLEEMHLLGSGVQSGSLAYVVPQAWIVSGNIRENIL	593
		: : : : : : : : : : :	
Db	503	VRKGILGICGNVSGSKSLLAALLGQMQLQKGVAVNGTLAYVVSQAWIFHGNVRENIL	562
QY	594	MGGAYDKARYLOVLHCCSLNRDLELLPFQDMTEIGERGLNLSGGQKQRIISLARAVYSDRQ	653
		: : : : : : : : : : : : : : : : :	
Db	563	FGKEYDHORYOHTVRVCGLOKDLNLPYGDLTEIGERGLNLSGGQQRISLARAVYSDRQ	622
QY	654	IYLLDDPLSAVDAAHVKGKHIFECIKKTLRGKTVVLVTHQLQYLEFCGQIILENGKICEN	713
		:	
Db	623	LYLLDDPLSAVDAAHVKGKHVFEECIKKTLRGKTVVLVTHQLQFLESCDEVILLEGEICEK	682
QY	714	GTHSELMQKKGKYAQLIQMH-----KEATSDMLQDTAKIAEKP--	752
		: :	
Db	683	GTHKELMEERGRYAKLIHNLRLGLQFKDPEHLNAAVMAFAPKESPAEREEDAGIIVLAPGN	742
QY	753	-KVESQALATSLAESLNGNAVPEHQLTQEEEMEEGSLSWRVYHHYIOAAGGYMVSCIIFF	811
		: : : : : : : :	
Db	743	EKDECKESETGSE--FVDTKVPEHQLIQTESPQEGTVTWKTYHTYIKASGGYLLSLFTVF	800
		: : : : : : : :	
QY	812	FVVLIVFLTIFSEFWLSYWLQGGSGTNSSRESNGTMADLGNADNPQLSFYQLVYGLNAL	871
		: : : : : : : :	
Db	801	LFLLMIGSAAFSNNWLGLDKSRMTCGPGQNRNRTMCEVCAVLADIGQHVYQWVYTASMV	860
QY	872	LLICVGVCSGGIFTKVTRKASTALHNKLFNKVFCPMSEFDTIPIGRLLNCFAGDLEQLD	931
		: : : : : :	
Db	861	FMLVFGVTKGFTKTTTLMASSSLHDTVFDKILKSPMSFFDTTPTGRLMNRFSKDMDELD	920

Db 143 VLIHQILOQTERTSKGVWVGIGLCIALFATEFTKVVFWALAWAINYRTAIRLKVALSTLV 202
QY 243 FEKLIQFKSVHITSGEAISFTTGDVNVLFEGVCYGPLVLITCASLVICSISSYFIIGYT 302
Db 203 FENLVSEKTLTHISVGEVLNILSSDSYSLFEAALFCPLPATIPIILMVFCAYAFILGPT 262
QY 303 AFIAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQRIRVTSEVLTCIKLIKMYTWKPPFA 362
Db 263 ALIGISVYVIFIPVQMFMAKLSAFRRSAILVTDKRVQTMNEFLTCIRLIKMYAWEKSEFT 322
QY 363 KIIEDLRRKERKILLEKCGLVQSLTSITLFIPTVATAVWVLJHTSLKLIKLTASMAFSMLA 422
Db 323 NTIQDIRRRERKILLEKAGVFQSGNSALAPIVSTIAIVLTLSCHILLRRKLTAPVAFSVIA 382
QY 423 SLNLLRLSVFFPIAVKGLTNSKSAVMRKFFFLQESPVFYVQTLQDPSKALVFEETLS 482
Db 383 MENVMKFSIALPFSIKAMAEANVSLRMKKILIDKSPSYITQPEDPDTVLLLANATLT 442
QY 483 W-----QOTCPGIVNG-----ALELERNGHASEGMRPRDALGPREEGNSLGPHELKINLV 533
Db 443 WEHEARQESTPKLQNKRLCKKQORSEAYSESRPPAKGATGPPEQSDSLKSVLHSISFV 502
QY 534 VSKGMMGLVCGTGTSGKSSLLSAILEEMHLEGVGVQGSLAYVPOQAWIVSGNIRENIL 593
Db 503 VRKGKILGICGVNGSGKSSLLAALLGOMQLQKGVVAVNGTLAYVSOQAWIFHGNVRENIL 562
QY 594 MGGAYDKARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKORISLARAVYSDRQ 653
Db 563 FGEKYDHQRYQHTVRVCGLQKDLNSLNPYGDLTEIGERGLNSGGQKORISLARAVYSDRQ 622
QY 654 IYLLDDPLSAVDAAHVKGKHIFEECIKKTLRGKTVVVLVTHQLQYLEFCGQIILLENGKICEN 713
Db 623 LYLLDDPLSAVDAAHVKGKHVFEECIKTLRGKTVVVLVTHQLQFLESCDEVILLEGEICEK 682
QY 714 GTHSELMQKKGYAQLIQKMH-----KEATSDMLQDTAKIAEKPKV 754
Db 683 GTHKELMEERGRYAKLIHNLRLGLQKDPPEHLYNAAMVEAFKESPAEREEDAVLAPGNEKD 742
QY 755 ESQALATSLEESLNGNAVPEHQLTQEEEMEEGSLSWRVYHHYIOAAGGYMVSCIIFFV 814
Db 743 EGKESSETGSE--FVDTKVPEHQLIQTESPQEGTVTKTYHTYIKASGGYLLSLFTVFLFL 800
QY 815 LIVFLTIFSFWLSYWLEQSGTNSRESNGTMDLGNADNPQLSFYQLVYGLNALLLI 874
Db 801 LMIGSAAFSNWGLWLDKGSRTMCGPQGNRTMCEVGAVLADIGQHVIYQRYVTASVMFML 860
QY 875 CVGVCSSGIFTKVTTRKASTALHNKLFNKVF 904
Db 861 VEGVTKGFTKTTLMASSSLHDTVFDKHF 890

RESULT 9

Q8ST87
ID Q8ST87 PRELIMINARY; PRT; 1334 AA.
AC Q8ST87;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ABC transporter ABCC.10 (Bile pigment transporter 1).
GN ABCC10.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Anjard C., Loomis W.F.;
RT "Evolution of the ABC transporters of Dictyostelium."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;

RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF474342; AAL85713.1; -
DR EMBL; AC116959; AAM08448.1; -
SQ SEQUENCE 1334 AA; 149288 MW; B29253642F2B32E8 CRC64;

Query Match 27.5%; Score 1946.5; DB 5; Length 1334;
Best Local Similarity 34.7%; Pred. No. 1.9e-131;
Matches 459; Conservative 251; Mismatches 514; Indels 99; Gaps 25;

QY 85 PLDNAGLFSYLTVSWLTPLMIQSLRSRLDENTIPPLSVHDASDKNVQRLHRLWEEVSRR 144
Db 30 PEENSNFLSNLTFSWADGFVIHCFRNVQLSHLWDLASYDKSEYLAKKIAKSWEIEIQK- 88
QY 145 GIEKASVLLVMLR-FQTRRLIFDALLGICFCIASVLGPILIIPKILEYSEEQLGNVVH-- 201
Db 89 --PKPSYLRAGFRAFGKLQLLSIFLYAISVGIOFV-GPEILGRMVTFFVESKLGSTEDP 145
QY 202 GVGLCFALFLSECVKSLSPSSSWI--INQRTAIRFRAAVSSFAFEKLIQPKSVIH--ITS 257
Db 146 NMGYVVALIMFGTAMIGSFCTYHANRISFRTGDRLSIIVLVYKKAIKLSNSARSDTSP 205
QY 258 GEAISFTTGDVNYLFE--GVC-YGPLVLITCASLVICSISSYFTIGYTAFAIACYLLVF 314
Db 206 GQIVNLMSNDAQRMVEVFGMFNNGALAL---PQIIICLALLYKKIKGWPTFVGLGLMLAAI 262
QY 315 PLAVFMTMAVKAQHHTSEVSDQRIRVTSEVLTCIKLIKMYTWKPFKAKIIEDLRRKERK 374
Db 263 PFNGMAAKKLTETRKYLVSLSDSRVKATNEILQAIIKIKYAWEDSFAKKVIHRNNEIK 322
QY 375 LLEKCGLVQSLTSITLFIPTVATAVWVLHTSLKCLKLTASMAFSLASLNLRLSVFFV 434
Db 323 LLFSYSRYETILIVIIISALPTAAAILVISSYYGHEKSLDASRIFSALSYLNLRLPLGFL 382
QY 435 PIAYKGLTNSKSAVMRKFFFLQESPVFYVQTLQDPS--KALVFEETLSWQOTCPGIVN 492
Db 383 PIIIALGIGIQMGIAGKRVTDLFLLE-MKDIQIDNPSPNGVYMKNSTTTWNKL----- 435
QY 493 GALELERNGHASEGMRPRDALGPREEGNSLGPHELKINLVSKGMMGLVCGTGTSGKSS 552
Db 436 -----KEDSFG-----LKNINFEATGTSLTMVVSGVSGKST 467
QY 553 LLSAILEEMHLEGVGVQGSLAYVPOQAWIVSGNIRENILMCGAYDKARYLQVLHCCSL 612
Db 468 LVQAMLGELEIIDGEIGIKGSIAVYVPOQAWIINATLKENIIFGKELDEERYQKVEVCAL 527
QY 613 NRDELLPFGDMTEIGERGLNSGGQKORISLARAVYSDRQIYLLDDPLSAVDAAHVKGHI 672
Db 528 KRDIELFPQDSVEIGERGINLSGGQKQVRSIARAVYSDADVILDDPLSAVDSHVGKHL 587
QY 673 FEECIKKTLRGKTVVVLVTHQLQYLEFCGQIILLENGKICENGTHSELMQKKGYAQLIQK 732
Db 588 FHKCFKGILSSKTVILVANQLNLYLPADNTVVLKSGEIVERGYVELINSKLEFSSILEK 647
QY 733 MHKEAT-----SDMLQDTAKIAEKPKVESQALATSLEESLNGNAVPEHQLTQEE 781
Db 648 YGVDENVISKKDDIDEDQDT---IEKVEIDLNKDEKSQPKSKSSNT--DGTLISEE 702
QY 782 EMEEGSLSWRVYHHYIOAAGGYMVSCIIFFVVLIVFL-----TIFSFWLSYWLEQSG 836
Db 703 ESEQGAVAGKVYWKYVTAGGG-----LLFLVSMIFFLLETGSKTFSDDWLSHWQTESSE 756
QY 837 TNSS----RESNGTMDLGNADNPQLSFYQLVYGLNALLLICVGCSSGIFTKVTTRKAS 892
Db 757 RMESILLGEEPTG-----LTDDQNGLGIY---IGL-GMAAVFISVCKNFIYEYSVYAS 805
QY 893 TALHNKLFNKVFCPMSFFDTIPIGRLLNCFAGDLEQLDQLPIFSEQFLVLSLMVIAVL 952
Db 806 RAIHHELFNALLKKPMYFFDQTPIGRINRFRTRDLGDIDNLIATISITFTLTMLTVIATI 865

Db	966	EHSRSPLYDHVSASLEGITTIHTFQSNRFLVLEVLKKHLDCNSGAIFMQSAMRWLAVWL	1025
QY	1051	EIMTNLVTLAVALFAFGISSTPYSFKVMVAVNIVLQASSFQATARIGLETEAQFTAVER	1110
Db	1026	DLVVVMTAIVALLTVMLTGTVSPADAGMAIAFAVQMGSIGFQFAVRTQTELEAKMTSVER	1085
QY	1111	ILQYMKMCVSEAPLH-MEGTSCPQGWPHQGEIIFQDYHMKYRDNTPTVLHGINTIRGHE	1169
Db	1086	VSYADNIPEDGEWNTROGLDISSWPANGQINFSEVNLRYRKSHPLALNDITFEIKGGE	1145
QY	1170	VVGIVRTGSGKSSLMALFRLVEPMAGRILIDGVDICSIGLEDRLSKLSVIPQDPVLLS	1229
Db	1146	KVGIIIRGTGSGKSSLANLIFRLYPVTNGTIYIDGVDIRTVGLVKLRRGISATAQDPSLFS	1205
QY	1230	GTIRENLDPDRHTDQOIWDALERTFLTKAISKFPKKLHTDVVENGNGNFSVGERQLLCIA	1289
Db	1206	GTVRENLDPSLEYSDSMIWEALEKCHLKTTLVQSLDKKLEADVSHGGNFSVGERQLFCLA	1265
QY	1290	RAVLNRNSKIILIDEATASIDMETDTLIQRTIREAFQGCCTVLVIAHRVTTVLNCDHILVMG	1349
Db	1266	RALLMKSRIVILDEATASVDAGTDKLIQEVIKTVFADATVIIIAHRLDNVRNMDRIMHLK	1325
QY	1350	NGKVVFEFDRPEVLRRKPGSLF	1370
Db	1326	NGKLINFTTPOEMFKDDWSVY	1346
RESULT 11			
Q8T6G8			
ID	Q8T6G8	PRELIMINARY;	PRT; 1264 AA.
AC	Q8T6G8;		
DT	01-JUN-2002 (TrEMBLrel. 21, Created)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	ABC transporter ABCC.12 (Fragment).		
GN	ABCC12.		
OS	Dictyostelium discoideum (Slime mold).		
OC	Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.		
OX	NCBI_TaxID=44689;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=AX4;		
RA	Anjard C., Loomis W.F.;		
RT	"Evolution of the ABC transporters of Dictyostelium.";		
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF474344; AAL85715.1; -		
FT	NON_TER 1264 1264		
SQ	SEQUENCE 1264 AA; 141643 MW; 3961B6A0AC66C1D6 CRC64;		
Query Match 27.1%; Score 1918; DB 5; Length 1264;			
Best Local Similarity 34.1%; Pred. No. 2e-129;			
Matches 450; Conservative 246; Mismatches 491; Indels 132; Gaps 26;			
QY	85	PLDNAGLFSYLTVSWLTPLMIQSLRSLRDENTIPPLSVHDASDKNVQRLHRLWEEVSRR	144
Db	34	PEENSNFLSNLTFSWADGFVIHCFRNVQLSHLWDLASYDKSEYLAKKIAKSWIEIQK-	92
QY	145	GIEKASVLLVMLRFQRTRLIFDALLGICFCI-----ASVLGPILIPKILEYSEE	194
Db	93	--PKPSYLRAGR-----AFGKLHCISLFFYSIYVGSQFVGPEILSRMVTFFVES	140
QY	195	QLGNVVH--GVGLCFALFLSECVKLSLFSWSSI--INQRTAIRFRAAVSSFAFEKLIQFK	250
Db	141	KLGTSTEDPNMGYYIALIMFGTAMIGSFCNCYQANRVTVRTGDRLRSIIIVLDVYKKAIKLS	200
QY	251	SVI--HITSGEAISFFTGDVNYLFEGVCYGPL--VLITCASLVICSISYFLIGYTAFIA	306
Db	201	NSARSNTSPQIVNLSNDAQRMIE--VFGLNGLFALPQIIICLALLYEKIGWPTFVG	258
QY	307	ILCYLLVFPFLAVFMRMAVKAQHHTSEV---SDQIRVTSEVLTCTIKLIKMYTWKPPFA	362
Db	259	LGLMLAAIP----FNLAAKKLTETRRILIGHTDGRVKVTSEILQAMKIIKLYAWEDSFA	314

QY	363	KIIEDLRRKERKLEKCGLVQSLTSITLFIPTVATAVWVLIHTSLKLKTASMAFSMLA	422
Db	315	KKVLDRRNNEIKLLFSFTRYRTILIAMIGAIPTAASILVFSTYYGYNGSLDACKIFSALS	374
QY	423	SNLLRLSVFFVPIAVKGLTNSKSAYMRFKKFFLQESPVFYVQTLQDPS--KALVFEAT	480
Db	375	YLNLLKIPGLPILIALGIOMQIAASKRVTDFLLLPE-MKEVQQIDNPSPNGVYMKNST	433
QY	481	LSWQQTCPGIVNGALELERNGHASEGMTRPDALGPEEENSLGPELHKINLVVSKGMML	540
Db	434	TTWN-----KEKEDSFG--LKNINFEAKGQSILT	459
QY	541	GVCGNTGSGKSSLLSAILEEMHLLGSGVGVQGSLAYVPPQAWIVSGNIRENITLMGGAYDK	600
Db	460	MVGVSGSGKSTLVQAMLGELETIDGEIGIKGSIAYVPPQAWIINATLKENIIFGKELDE	519
QY	601	ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKQRISLARAVYSDROIYLLDDP	660
Db	520	ERYQKVLEVCAKRDIELFPQGDSVEIGERGINLSGGQKQRVSIARAVYSDADVYILDDP	579
QY	661	LSAVDAHVGKHIFEECIKKTLRGKTVVLVTHQLQYLEFCGQIILLENKICENGTHSELM	720
Db	580	LSAVDSHVGHLFHKCFKGILSSKTVILVANQINYLFPADNTVVLKSGEIVERGTYYELI	639
QY	721	QKKGKYAQLIQKM-----HKEATSDMLQDTAK--IAEKPKVESQALAT'SLEESLNG	769
Db	640	NAKLEFASLLQEYGVNDENTKGDSDDDDKDDKKKEKVEKPKQSDK-----	687
QY	770	NAVPEHQLQTEEMEEGSLSWRVYHHYIOAAGYMWVSCIIFFFVVLIVFL----TIFSFW	825
Db	688	----DGTLISEEEAQGAVAGKYVYTAGGG-----LLFLFAMILFLLETGSKTFTDW	738
QY	826	WLSYWLEQGSNTSS---RESNGTMADLGNADNPQLSFYQLVYGLNALLLICVGVCSS	881
Db	739	WLSHWQTESSERMESILLGEPTG-----LTDDQNLGIY-IGVGMASIIIVTVVRTFS-	789
QY	882	GIFTKVTRKASTALHNKLFNKVFCPMSFFDTPIGRLLNCFAGDLEQLDQLLPFSEQF	941
Db	790	--FFEYAVRAAHSIHHELFNALLKKPMSFFDQTPGLRIINCFTRDLDIIDNLIATSIQAF	847
QY	942	LVLSLMVIALLIVSVLSPYILLMGAIIIVICFIYYMMFKKAIGVFKRLENYSRSPLFSH	1001
Db	848	FTLMLSVLATLILISIIYPWLLIPLAPICILFFILOFYRYTYSRGLQRIEAITRSPIFNH	907
QY	1002	ILNSLQGLSSIHYGK-TEDFISQFKRLTDAQNNYLLLFLSSTRWMALRLEIMTNLVTLA	1060
Db	908	FSETLNGVVSIRAYKKQKENILKNQKRLDDNNNCYLT-L-QAMNRWLGRDLFLGNLIVFF	966
QY	1061	VALFVAFGISSTPYSFKVMVAVNIVLQASSPQATARIGLETEAQFTAVERILQYMKMCVS	1120
Db	967	SCIFITLKKDTISPSDVGVLVSYALSITSNLNQGVLQAADTETKMNSVERISQYIRGAV-	1025
QY	1121	EAPLHMEGTSQPGWPQHGEIIFQDYHMKYRDNTPTVLHGINTIRGHEVVGVIVRTGSG	1180
Db	1026	EAPQIIDDCRPSDPWPINGSIKFDNLVMRYREGLDPVLKGITCEIKAKEKIGIVRTGAG	1085
QY	1181	KSSLGMALFRLVEPMAGRILIDGVDICSIGLEDLRSLKSVIPQDPVLLSGTIRFNLDPPD	1240
Db	1086	KSSIVLALFRLIEASEGSISIDGENIAKFLKDLRRNLAIIPQDPVPLFSGTLRENLDPFN	1145
QY	1241	RHTDQOIWDALERTFLTKAISKFPKKLHTDVVENGNGNFSVGERQLLCIARAVLRNSKIIL	1300
Db	1146	ECPDHELWSILDDIQLSKVFKSTE EGLNSKVTENGENFSGVQORQLIVLARALLRKPILV	1205
QY	1301	IDEATASIDMETDTLIQRTIREAFQGCCTVLVIAHRVTTVLNCDHILVMGNGKVVFEFDRP	1359
Db	1206	LDEATASVDGQSDSLIQATIRNKFSNCTILTIAHRLNTIMDSKIMVLDAGKISEFDEP	1264

RESULT 12

Q8T1F1

ID Q8T1F1

PRELIMINARY;

PRT; 1323 AA.

Db 1 WLNPLFSIGSKRRLEEDDMFNVLPEDRSKKLGEELOQSYWDQEKKEKAAKELKTPKL----- 55

QY 159 QRTRLIFD-----ALLGICFCI---ASVLGPILIIIPKILEYSEELGNVNVHGVGLCFAL 209

Db 56 --TKAIIRCYKWSYAVLGVFTLIEESIKVIQPV-FLGKLIKIFYENRHDDMAALSEAYGY 112

QY 210 FLSECVKSLSPSSSWIIN-----QRTAIRFRAAVSSFAFEKLIQPKSVI--HITSGEA 260

Db 113 ATGVCLSTLGLA--LLHLYFYHVORAGMKIRIAMCHMIYRKALCLSAAMGQTTTGQI 169

QY 261 ISFFTGDVNYLFEQVCY-----GPLVLITCASLVICSISSYFIIGYTAFAIAILCYLLVF 314

Db 170 VNLLSNDVNKFEDELTIHFLVWVGPLQAAAIVIGLLWQEIGPSCLAG----MAVLVFLM-- 223

QY 315 PLAVFMTRMVAKQAHHITSEVSDQRIRVITSEVLTICIKLIKMYTWKPKFAKIIEDLRRKERK 374

Db 224 PLQTMFGKLFYSKYSKTAALTDSRIRTMNEVSVGIRIIKMYAWEKPFAMLVNDVRRKEIS 283

QY 375 LLEKCGLVQSITSITLF-----IIPTVATAVWVLIHTSLKCLKLTASMAFSMLASNLRLLS 430

Db 284 KIMSSSYLRGLNMAEFFTANKIILFVTFVYVLVGNT---MSASRVFVAVSLYSAVRLT 339

QY 431 V-FFVPIAVKGLTNSKSAVMRFKKFFLOESPV-FYVQTLQDPSK--ALVFEETLSWQQT 486

Db 340 VTLFFPAAIEKVSESAISRRIKKFKLLDDELVKNHLPLSQEEKKEPSVEMQDLICYWDKT 399

QY 487 CPGIVNGALELERNGHASEGMTRPDRDALGP EEGNSLGPHELKINLVVSKGMMLVGCGNT 546

Db 400 L-----DA-----PTLQNVCFVTKPGQLLAVIGPV 424

QY 547 GSGKSSLLSAILEEMHLLLEGVGVQCSLAYVVPQAAWIVSGNIRENIIILMGGAYDKARYLQV 606

Db 425 GAGKSSLLSTVLGELPAEKGVIKVKGELTYASQQPWFPGTIRSNILFGKELQPORYERV 484

QY 607 LHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRI SARAVYSDROIYLLDDPLSAVDA 666

Db 485 LRACALKRDMELLPDGDLTIVIGDRGATLSGGQKQARVN LARAVYQADAIYLLDDPLSAVDA 544

QY 667 HVGKHIFEICKTKTLRGKTVVLVTHQLQYLEFCGQIILLENGKICENGTHSELMQKKGY 726

Db 545 EVSRHLFEQCVCGILKDKPRILVTHQLQYLKAANQILVLKEGHMVARGSYSELOQSGLDF 604

QY 727 AQLIQMKHKEATSDMLQDTAKIAEKPV-----ESQALATSLEESLNGNAVPEHQLTQ 779

Db 605 TSLLKKDEEBESGSEKGEAPRSPRSTVSONSVRSHSSSVLSVKRDSQDQLPAEPVHTMA- 663

QY 780 EEEMEEGSLSWRVYHHY IQAAGGYMVSCIIFFVVLIVFLTIFSPWMLS YW-LEQGS GTN 838

Db 664 EESRSEGNIGIRMYWKYFRAGANVVMVLVLLNLLAQTFFI LDQWMLS YWATEQEKL DH 723

QY 839 SSRESNGTMADLGNADNPQLSFYQLVY-GLNALLLICVGCSSGIFTKVTRKASTALHN 897

Db 724 NTNNTNTNTSAGNTTQQLDLNLFYLG IYAGLTGATIVFGEMRCLIMENALVSSAET-LHN 782

QY 898 KLFNKVFCRPMSSFDDTIPIGRLLNCFAGDLEQLDQLLP I FSEQLVLSLMVIAVLLIVSV 957

Db 783 RMFNSILRTPVRFFDINPIGRILNRF SKDIGHLD SLLPWFVDFIQVFLQIVGVIAVASS 842

QY 958 LSPYILLMGAIIMVICFI----YVMFPKKAIGVFKRL ENYSRSPLEF SHILNSLQGLSSIH 1013

Db 843 VIPWILI-PVLP LLI C FLRRYFLRTRSDV---KRIESTTRSPV FSHLSSSLQGLW TIR 898

QY 1014 VYGKTEDFISQFKRLTDAQNNYLLLF LSSTRWMALRLEIMTNLVTLAVALFVAFGISSTP 1073

Db 899 AFKAEEERFOQTEDAHQDLHSEAWFLFTLSRWF AVRLDGMCS-VFV TITAFGCLLLKDTM 957

QY 1074 YSFKV-MAVNIVLQLASSFQATARIGLETEAQFTAVERILQYMKMCVSEAPLHMEGTS CP 1132

Db 958 NAGDVGLALSYAVTLMGMFQWVGRQSAEVENMMT SVERVVEYTEL-ESEAPWETQKRPS P 1016

QY 1133 QGWPQHGEIIFQDYHMKYRONTPTVHLGINLTIRGHEVVGVIGRTSGSKSSLGMA LFR LV 1192

Db 1017 D-WPNRGLITFDRVNF SYSSDGPVVLKNISAMFRPREKVGIVGTGAGKSSLSALFRLS 1075

QY 1193 EPMAGRILIDGVDICSIGLEDLRSLKLSVIPQDPVLLSGTIRFNLDPFDRHTDQQIWDAL E 1252

Db 1076 EP-EGKILVDGVLTSEIGLHDLRQKMSLIPODPVLFTGTMRKNLDPFNQHSDDLWKAL E 1134

QY 1253 RTFLTKAISKFPKKLHTDVVENGGNFSVGERQLLCIARAVLRNSKIILIDEATASIDMET 1312

Db 1135 EVQLKAAVEELPGKLETELAESGNSFSVGRQLVCLARAILRKNRVLIIDEATANVDPR T 1194

QY 1313 DTLJORTIREAFQGCCTVLVIAHRVTTVLNCDHILVMGNGKVVEFDRPEVLRKKPGSLFAA 1372

Db 1195 DELIQKTIIRDKFEKCTVLTJIAHRLNTIIDSRLILVDAGRIHEYDAPHVLLQNQSGIFYK 1254

QY 1373 LM-----ATATSSLR 1382

Db 1255 MVQQTGKA EATSLLO 1269

RESULT 14

Q8T6H5

ID Q8T6H5 PRELIMINARY; PRT; 1247 AA.

AC Q8T6H5;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE ABC transporter ABCC4

GN ABCC4.

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

OX NCBI_TaxID=44689;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AX4;

RA Anjard C., Loomis W.F.;

RT "Evolution of the ABC transporters of Dictyostelium.";

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF474336; AAL85707.1; -

FT NON_TER 1247 1247

SQ SEQUENCE 1247 AA; 139561 MW; 4926820E45C7519C CRC64;

Query Match 26.9%; Score 1903.5; DB 5; Length 1247;

Best Local Similarity 34.4%; Pred. No. 2.le-128;

Matches 448; Conservative 247; Mismatches 494; Indels 115; Gaps 26;

QY 85 PLDNAGLFSYLTVSWLTPLMIQSLRSLRDENTIPPLSVHDASDKNVQRLHRLWEEEVSR R 144

Db 30 PEENSNFLSNLTFESWADGFVIHCFRNVQLSHLWDLASYDKSEYLAKKIAKSWEIEIQK- 88

QY 145 GIEKASVLLVMLR-FQTRRLIFDALLGICFCIASVLGPILIIPKILEYSEEQLGNNVH-- 201

Db 89 --PKPSYLRAGFRAFGKQLQLLSIFLYAISVGIOFV-GPEILGRMVTFFVESKLGSTEDP 145

QY 202 GVGLCFALFLSECVKSLSPSSSWIINORTAIRFRAAVSSFAFEKLIQFKSVIHITS-GEA 260

Db 146 NMGYYYALIMF---GTAMIGSFCTYHANRISFRTG-----DPIKLSNSARSDTSPGOI 195

QY 261 ISFFTGDVNYLFE--GVC-YGPLVLITCASLVICSISSYFIIGYTAFAIILCYLLVFPLA 317

Db 196 VNLMNSDAQRMVEVFGMFNNGALAL--PQIIICLALLYKKIGWPTFVGLGLMLAAIPFN 252

QY 318 VFMTRMAVKAQHHTSEVSDQRIRVITSEVLTICIKLIKMYTWKPKFAKIIEDLRRKERKLL E 377

Db 253 GMAAKLTETRKYLVSLSDSRVKATNEILQAIKIIKLYAWEDSPAKKVI EHRNNEIKL LF 312

QY 378 KCGLVQSLSITLFIIPTVATAVWVLIHTSLKCLKTASMAFSMLASNLRLSLVFFVPIA 437

Db 313 SYSRYRTILVIISALPTAAAILVISSYGYGHEKSLDASRIFSALSYNLRLPLGLFLPII 372

QY 438 VKGLTNSKSAVMRKKFFFLQESPVFVQTLQDPS--KALVFEETATLSWQQTCPGIVNGAL 495

Db 373 IALGIQMGIAGKRVTDFLLLP E-MKDIQQIDNPSPLPNGVYMKNSTTTWNKL----- 422

QY 496 ELERNGHASEGMTRPDRDALGP EEEGNSLGP ELHKINLVVSKGMMLGVCGNTGSGKSSLLS 555

Db 1066 YLSIY---VAFSCGTIAATFLRSFSMVFGSI--KSKLFHEKMFKAIVILSPMSFFDTT 1119
QY 915 PIGRLNCFAGDLEQLDQLLPFSEQFLVLSLVMVIAVLIVSVLSPYILLMGAIIIMVIFC 974
Db 1120 PIGRILNRFKSKDQLTIDESIARTLGMFLNTFCQVVGSIIVIAWVSPFIIL--AMVPVGAL 1177
QY 975 IYMM--FKKAIGVFKRLENYSRSPFLFSHILNSLQGLSSIHYGKTEDFSQFKRLTDAQ 1032
Db 1178 FYFIQKYLYLNSRELTRLEGVSRSPYIAHFESETLAGVTTIRAFKDVAREFTENERLLDEN 1237
QY 1033 NNYLLFLSSTRWMALRLEIMTNLVTLAVAFVAFGISSTPYSFVKVMVAVNIVLQLASSFQ 1092
Db 1238 QKCYINISSNRWLAIREFLGACLVSCAVLYTVIARSRIEAGTAGLVITYALAITGMN 1297
QY 1093 ATARIGLETEAQFTAVERILQYMKMCVSEAPLHMEGTSCPQGWPHQGEIIFQDYHMKYRD 1152
Db 1298 WMVRMSCDLENSVVSIERIQEYC-LLPSKAPL-FNDKSVPMSPWPSHGKIVFKNLWLTyre 1355
QY 1153 NTPTVLHGINLTIRGHEVVGIVGRTGSGKSSSLGMALFRLVEPMAGRILIDGVDICSIGLE 1212
Db 1356 GLDPVLRGINCTIEPKTKVGIVGRTGAGKSSLTQALFRLVEPLRGTTIEDGIDITELGLN 1415
QY 1213 DLRSKLSVIPQDPVLLSGTIRFNLDPFDRHTDQOIWDALERTFLTKAISKFPKLLHTDVV 1272
Db 1416 PLRSRMAIIPQDPVLFAGSVRYNLDPFQDYDDHEIWEAIEAENAHLLKAIKDLGGGLDAMVQ 1475
QY 1273 ENGNFSVGERQLLCIARAVLRNSKIIILIDEATASIDMETDTLIQRTIREAFQGCOTVVI 1332
Db 1476 DGGDNFSVGQRQLLVIGRALLKKANIIVLDEASSIDIASDALIQETIRTKFADCTVLT 1535
QY 1333 AHRVTTVLNCDHILVMGNGKVVEFDRPEVLRKKPGSLFAALMATATS 1379
Db 1536 AHRGLTIADSKIMVLDKGELIEYDSPSELLKNQDSIYYSLVKASES 1582

Search completed: July 21, 2003, 09:41:22
Job time : 118 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2003, 09:29:35 ; Search time 29 Seconds
(without alignments)
1976.562 Million cell updates/sec

Title: US-10-087-782A-31
Perfect score: 7071
Sequence: 1 MTRKRTYWVPNSSGGLVNRG.....RKKPGSLFAALMATATSSLR 1382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2694	38.1	1436	1 MRP5_RAT	Q9qym0 rattus norv
2	2683	37.9	1436	1 MRP5_MOUSE	Q9rlx5 mus musculus
3	2665.5	37.7	1437	1 MRP5_HUMAN	O15440 homo sapien
4	1890	26.7	1325	1 MRP4_HUMAN	O15439 homo sapien
5	1855.5	26.2	1522	1 MRP3_RAT	O88563 rattus norv
6	1848.5	26.1	1527	1 MRP3_HUMAN	O15438 homo sapien
7	1818	25.7	1564	1 MRP2_RABIT	Q28689 oryctolagus
8	1810.5	25.6	1541	1 MRP2_RAT	Q63120 rattus norv
9	1767	25.0	1531	1 MRP1_HUMAN	P33527 homo sapien
10	1766.5	25.0	1515	1 YCFI_YEAST	P39109 saccharomyc
11	1761	24.9	1545	1 MRP2_HUMAN	Q92887 homo sapien
12	1734.5	24.5	1478	1 YAWB_SCHPO	Q10185 schizosacch
13	1620.5	22.9	1477	1 YOR1_YEAST	P53049 saccharomyc
14	1585	22.4	1592	1 YHD5_YEAST	P38735 saccharomyc
15	1572	22.2	1580	1 ACC8_HUMAN	Q09428 homo sapien
16	1565	22.1	1503	1 MRP6_HUMAN	O95255 homo sapien
17	1548	21.9	1661	1 YBT1_YEAST	P32386 saccharomyc
18	1542.5	21.8	1581	1 ACC8_RAT	Q09429 rattus norv
19	1537.5	21.7	1581	1 ACC8_CRICR	Q09427 cricetus cr
20	1536	21.7	1559	1 BPT1_YEAST	P14772 saccharomyc
21	1534.5	21.7	1546	1 ACC9_MOUSE	P70170 mus musculus
22	1528	21.6	1545	1 ACC9_RAT	Q63563 rattus norv
23	1517	21.5	1549	1 ACC9_HUMAN	O60706 homo sapien
24	1513	21.4	1549	1 ACC9_RABIT	P82451 oryctolagus
25	1492	21.1	1502	1 MRP6_RAT	O88269 rattus norv
26	1397.5	19.8	1492	1 CFTR_SQUAC	P26362 squalus aca
27	1388.5	19.6	1548	1 MDR_LEITA	P21441 leishmania
28	1372	19.4	1481	1 CFTR_BOVIN	P35071 bos taurus
29	1371.5	19.4	1480	1 CFTR_HUMAN	P13569 homo sapien
30	1370	19.4	1481	1 CFTR_SHEEP	Q00555 ovis aries
31	1367.5	19.3	1476	1 CFTR_MOUSE	P26361 mus musculus
32	1338.5	18.9	1485	1 CFTR_XENLA	P26363 xenopus lae
33	1332	18.8	1427	1 ABC1_SCHPO	Q92337 schizosacch

ALIGNMENTS

RESULT 1				
MRP5_RAT				
ID	MRP5_RAT	STANDARD;	PRT;	1436 AA.
AC	Q9QYM0;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Multidrug resistance-associated protein 5.			
GN	ABCC5 OR MRP5.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;			
RA	Homma M., Suzuki H., Sugiyama Y.;			
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: ACTS AS A MULTISPECIFIC ORGANIC ANION PUMP WHICH CAN			
CC	TRANSPORT NUCLEOTIDE ANALOGS (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB020209; BAA88897.1; -.			
DR	HSSP; P13569; INBD.			
DR	InterPro; IPR003593; AAA_ATPase.			
DR	InterPro; IPR003439; ABC_transportr.			
DR	InterPro; IPR001140; ABCtranprtTM.			
DR	Pfam; PF00005; ABC_tran; 2.			
DR	Pfam; PF00664; ABC_membrane; 2.			
DR	ProDom; PD000006; ABC_transportr; 2.			
DR	SMART; SM00382; AAA; 2.			
DR	PROSITE; PS00211; ABC_TRANSPORTER; 2.			
KW	ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.			
FT	TRANSMEM 179 199 POTENTIAL.			
FT	TRANSMEM 219 239 POTENTIAL.			
FT	TRANSMEM 296 316 POTENTIAL.			
FT	TRANSMEM 317 337 POTENTIAL.			
FT	TRANSMEM 400 420 POTENTIAL.			
FT	TRANSMEM 426 446 POTENTIAL.			
FT	TRANSMEM 608 628 POTENTIAL.			
FT	TRANSMEM 847 867 POTENTIAL.			
FT	TRANSMEM 916 936 POTENTIAL.			
FT	TRANSMEM 996 1016 POTENTIAL.			
FT	TRANSMEM 1017 1037 POTENTIAL.			
FT	TRANSMEM 1101 1121 POTENTIAL.			
FT	TRANSMEM 1126 1146 POTENTIAL.			
FT	NP_BIND 595 602 ATP (POTENTIAL).			

Q00554 oryctolagus
P36028 saccharomyc
P08183 homo sapien
P34712 caenorhabdi
P06795 mus musculu
P21449 cricetus
P21448 cricetus
Q00449 drosophila
P43245 rattus norv
P21447 mus musculu
P23174 cricetus
O95342 homo sapien

FT	NP_BIND	1226	1233	ATP (POTENTIAL).
FT	CARBOHYD	494	494	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	636	636	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	684	684	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	889	889	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	896	896	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1043	1043	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1328	1328	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1416	1416	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1436 AA; 160855 MW; 10FE53B800531598 CRC64;		
Query Match 38.1%; Score 2694; DB 1; Length 1436;				
Best-Local Similarity 41.0%; Pred. No. 3e-165;				
Matches 559; Conservative 269; Mismatches 481; Indels 56; Gaps 10;				
QY	62	GYDAALRTMIPRPKPRFPAPQPLDNAGLFSYLTVSWLTPL-MIQSLRSRLDENTIPPL	120	
Db	79	GKYHHGLSALKPFRTTTKH--QHPVDNAGLFSYMTFSWLSPLAQVHVHKKGELLIMEDVWPL	136	
QY	121	SVHDASDKNVQRLHRLWEEVSRRGIEKASVLLVLMRFQRTLRIFDALLGICFCIASVLG	180	
Db	137	SKYESSDVNCRRLERLWQEEELNEVGPDAASLRRVWVIFCRTLILSIVCLMITQLAGFSG	196	
QY	181	PILIIPKILEYSEEOLGNVHVGVGLCFALFLSECVKLSLFSFSSWIINQRTAIRFAAVSS	240	
Db	197	PAFVVVXHLLEYQTATESNLQYSLLLVLGLLLETVVRSWSLALTWALNYRTGVRLRGAVLT	256	
QY	241	FAFEKLIQFKSVIHITSGEAI5FFTGVDVNVLFEGVCYGPLVLITCASLVICSIS5YFIIG	300	
Db	257	MAFKKILKLNKEKSLGELINICSNDGQRMFEAAAAGSLLAGGPVVAILGMIYNVILG	316	
QY	301	YTAFIAILCYLLVFPPLAVFMTMRMAVKAQHHTSEVSDQRIRVTSEVLTCTIKLIKMYTWEKP	360	
Db	317	PTGFLGSAVFILFYPMAMFVSRLTAYFRRKCVCAATDDRQVKMNEVLTYYIKFIKMYAWVKA	376	
QY	361	FAKIIEDLRRKERKLLKCGLVQSLTSITLFIPTVATAVAVVLIHTSLKLTASMAFSM	420	
Db	377	FSQCQVKIREEERILEKAGYFQ5ITVGVAPIVVVIVASVVTF5VHMTLGFDLTAAQAFTV	436	
QY	421	LASNLRLSVFFVPIAVKGLTNSKSAVMRFFKFFLQESPVFYVQTLQDPSKALVFEAT	480	
Db	437	VTVFN5MTFALKVTPF5VKLSLSEASVAVDRFKSLFLMEEVHMIKNKPASPHIKIEMKNAT	496	
QY	481	LSWQ-----QTCPCGIV-----NGALELERNGHA---SEG	507	
Db	497	LAWDSSHSSTQSSPKLTPKVKKDKRAPKKGKESRQLQHTHQAVLAEQKGHLLDSDER	556	
QY	508	TRPRDALGP EEGNS--LGPELHKINLV5KGMMLGVCNGTSGKSSLLSAIL EEMHLL E	565	
Db	557	PSPEEEGKQIHAGSMRLQRTLNYIDLEIEEGKLVGICG5VSGKTSLSAILGQMTLLE	616	
QY	566	GSVG5QSLAYVPPQAWIVSGNIRENIMLGGAYDKARYQLVLHCCSLNRDLELLPFGDMT	625	
Db	617	G5IAVSGTFAYVAQAWILNATLRDNILFGKEFDEERYNSVLNSCCLRPDLAILPNSDLT	676	
QY	626	EIGERGLNSGGQKORISLARAVYSDRQIYLLDDPLSAVDAAHVKGKHIFEECIKKTLRGKT	685	
Db	677	EIGERGANL5GGQRISLARALYSDRSIYILDDPLSALDAHVGNHIFNSAIRKRLKSKT	736	
QY	686	VVLVTHQLQYLEFCGOIILLENGKICENGTH5ELMQKKGKYAQLIQMHKEATSDMLQDT	745	
Db	737	VLFVTHQLQYLVDCDEVIFMKEGCITERGTH5EELMNLNGDYATIFNNL-----LLGET	789	
QY	746	AKIAEKPKVESQALATSLEESLNGNAV-----PEHQLTOEEEMEEGSL5WRVYHHYI	797	
Db	790	PPVEINSKKEASGQK5QDKGPKPGSVKKEKAVK5EEGQLVQVEEKGG5VP5W5VWYI	849	
QY	798	QAAGGYMV5CIIFF5FVVLIVELTIF5FWL5Y5WLEQSGTNS5RESN-GTMADLGNADN	856	
Db	850	QAAGG5PLAFLVIMVLFMLNVG5TAF5TWL5Y5WIKQ5G5GN5TVFEGNR55VSD--5MRDN	907	
QY	857	PQLSFYQLVYGLNALLLICV5G55GIFTK5VTRKASTALHNKLFN5K5VRC5P5M5FFDTIPI	916	

Db	908	PFLQYYASIYALSMAVMLILKAIRGVVVFVKGTLRASSRLHDELFRRLILRSPMKFFDTPPT	967	
QY	917	GRLNCFAGDLEQLDQLLP5IF5EQFLVLSL5M5VIAVL5V5L5P5YILM5CAIIMVICFIY	976	
Db	968	GRILNRF5KDMDEV5RLP5FQAEMFIQN5VIL5V5FCVGM5IAG5VFP5WFL5VAVG5PLL5L5FSVL	1027	
QY	977	YMF5KKAIG5VFKRLE5N5SR5PL5F5SHIL5N5I5Q5L5SS5IH5V5Y5G5K5TED5FI5Q5K5RL5T5DAQN5YL	1036	
Db	1028	HIV5RVL5IREL5KRL5DNIT5Q5P5FL5SHIT5SS5T5Q5GLAT5IHAYN5K5RQ5EFL5H5RYQ5EL5LD5D5NQAPF	1087	
QY	1037	LLFL5S5TR5W5MAL5LEIM5TN5L5VTL5A5V5AL5F5AG5IS5T5P5Y5FK5VM5AVN5IVL5Q5LAS5F5Q5ATAR	1096	
Db	1088	FLFT5CAM5RL5AV5RL5DL5L5IS5IAL5ITTT5GL5M5IVL5MHG5I5P5SAY5AG5LAIS5AV5QLT5GL5FO5FTVR	1147	
QY	1097	IGLE5TEA5Q5T5A5VER5IL5Q5Y5MK5MC5V5EAPL5H5ME5GT5SCP5Q5GW5P5QH5GEI5IFQ5DY5H5MK5Y5RD5NTPT	1156	
Db	1148	LASE5TEA5R5FT5S5VER5INH5YIKT5LSLE5AP5ARI5KN5KAP5PH5DW5PQ5E5GEIT5E5NA5EM5RY5REN5LPL	1207	
QY	1157	VLHG5INL5TIR5GHE5V5GIV5GRT5G5G5K5SL5G5MAL5F5RL5VE5PM5AGR5IL5ID5G5VD5ICS5IG5LED5LRS	1216	
Db	1208	VLK5K5V5FT5IK5PK5E5IG5IV5GRT5G5G5K5SL5G5MAL5F5RL5VEL5SG5CI5KID5GV5R5IS5DIG5LAD5LRS	1267	
QY	1217	KLSV5IPO5QDP5VLL5SGT5IR5NL5DP5DR5HT5DQI5WD5ALERT5LT5K5AIS5K5PK5KL5HT5DV5V5ENG5G	1276	
Db	1268	KLTI5I5PQ5E5PV5L5SGT5VR5NL5DP5FN5Q5YTE5EQI5WD5ALERT5HM5KE5CIA5QL5PL5K5LE5SE5VM5ENG5D	1327	
QY	1277	NFSV5GER5QL5CI5AR5AV5L5R5NS5KI5L5IDE5AT5ASID5MET5DTL5I5Q5RT5IRE5AF5Q5CT5VL5V5IAHRV	1336	
Db	1328	NFSV5GER5QL5CI5AR5ALL5RH5CK5IL5L5DE5ATA5MD5T5ET5DL5LI5Q5ET5IRE5AF5AD5CT5ML5T5IAHRL	1387	
QY	1337	TTVL5NCD5HIL5VM5G5NK5V5VE5F5DR5PE5VL5RK5KP5G5SL5FAAL5MAT5AT5SSL	1381	
Db	1388	HTVL5G5DR5IM5VLA5QG5QV5VE5F5D5TP5SVLL5S5ND5SR5FY5AM5CAA5AENKV	1432	
RESULT 2				
MRP5_MOUSE STANDARD; PRT; 1436 AA.				
ID	MRP5_MOUSE			
AC	Q9RLX5; O88284;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Multidrug resistance-associated protein 5 (ABC transporter MOAT-C) (SMRP).			
GN	ABCC5 OR ABCC5A OR MRP5.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=20184734; PubMed=10721709;			
RA	Suzuki T., Sasaki H., Kuh H.J., Agui M., Tatsumi Y., Tanabe S.,			
RA	Terada M., Saijo N., Nishio K.;			
RT	"Detailed structural analysis on both human MRP5 and mouse mrp5 transcripts.";			
RL	Gene 242:167-173(2000).			
RN	[2]			
RP	SEQUENCE OF 1302-1436 FROM N.A.			
RA	Suzuki T., Kuh H., Nishio K.;			
RT	"Molecular cloning of mouse homologue of SMRP/MRP5.";			
RL	Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: ACTS AS A MULTISPECIFIC ORGANIC ANION PUMP WHICH CAN TRANSPORT NUCLEOTIDE ANALOGS (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial			

QY	419	SMLASLNLLRLSVFFVPIAVKGLTNSKSAMRFKFFFLQESPVFYVQTLQDPSKALVEEE	478
Db	435	TVVTVFNMTFALKVTPFSVKLSLSEASVAVDRFKSLFLMEEVHMIKNKPASPHIKIEMKN	494
QY	479	ATLSWQQTCPGIVNG-----	507
Db	495	ATLAWSSSHSIQNSPKLTPKMKDKRASRGKKEKVRQLQRTHEQAVLAEQKGHLDDSD	554
QY	508	TRPRDALGPEEEGNS-----LGPELHKINLVSKGMLGVCGNTSGSKSSLLSAILLE	559
Db	555	ERP----SPEEEEGKHIHLGHLRLQRTLHSIDLEIQEGLVGCISVSGSKTSLISAILG	610
QY	560	EMHLLGSGVGVQSLAYVPOQAWIVSGNIRENILMGGAYDKARYLQVLHCCSLNRDLELL	619
Db	611	QMTLLEGSIAISGTFAYVAQQAUILNATLRDNILFGKEYDEERYNSVLNSCCLRPDLAIL	670
QY	620	PFGDMTEIGERGLNLGGQKQRIISLARAVYSDRQIYLLDDPLSAVDHAVGKHIFEECIKK	679
Db	671	PSSDLTEIGERGANLGGGQQRORISLARALYSDRSIYILDDPLSALDAHVGNHIFNSAIRK	730
QY	680	TLRGKTVVLVTHOLQYLEFCGQIILLENGKICENGTHSELQKKGKYAQLIQKMHKEATS	739
Db	731	HLKSKTVLVFVTHOLQYLVDCDEVIFPMKEGCITERGTHEELMNLNGDYATIFNNLLGETP	790
QY	740	DMLQDTAKIAEKPKVESQALATSLEESLNGNAV--PEHQLTQEEEMEEGSLSRVYHHYI	797
Db	791	PVEINSKKETSGSOKKQDKGPKTGSVKKEKAVKPEEGQLVQLEEKQGSVPWSVGYVI	850
QY	798	QAAGGYMVSCIIFFVVLIVFLTFPSFWLSYWLQEGSGTSSRESNGTMADLGNADNP	857
Db	851	QAAGGPLAFLVIMALFMLNVGSTAFSTWLSYWIQSGNVTVTRGNETSVS-DSMKDNP	909
QY	858	QLSFYQLVYGLNALLLICVGVSSGIFTKVTRKASTALHNKLFNKVFCRCPMSFFDTPIG	917
Db	910	HMQYASIIYALSMAVMLILKAIRGVVFKGTLRASSRLHDLFRILRSPMKFFDTPTG	969
QY	918	RLLNCFAGDLEQLDLLPIFSEQFLVLSLMVIAVLIVSVLSPIYLLMGAIIMVICFIYY	977
Db	970	RILNRFSKDMDEVDRLPFQAEMFIQNVILVFFCVGMIAGVPFWFLVAVGPLVILFSVLH	1029
QY	978	MMFKKAIGVFKRLENYSRSPLESHILNSLQGLSSIHVYKGTEDFISQFKRLTDAQNNVLL	1037
Db	1030	IVSRVLIRELKRLDNITQSPFLSHITSSIOGLATIHAYNKGQEFHLRYQOELLDDNQAPFF	1089
QY	1038	LFLSTRWMALRLEIMTNLVTLAVALEFVAFGISSTPYSFKVMANVIVLQLASSFOATARI	1097
Db	1090	LFTCAMRWLAVRLDLISIALITTTGLMIVLMHGQIPPAYAGLAISYAVQLTGLFQFTVRL	1149
QY	1098	GLETEAQFTAVERILOYMKMCVSEAPLHMEGTSCPQGWPOHGEIIFQDYHMKYRDNPTTV	1157
Db	1150	ASETEARFTSVERINHYIKTLSLEAPARIKNKAPSPDPQGEVTFENAEAMYRENPLPV	1209
QY	1158	LHGINTLTIRGHEVVGIVRTGSGKSSGLMALFRLVEPMAGRILIDGVDICSIGLEDLRSK	1217
Db	1210	LKKVSFTIKPKKEKIGIVRTGSGKSSGLMALFRLVELSGGCIKIDGVRISDIGLADLRSK	1269
QY	1218	LSVIPQDPVLLSGTIRFNLDPRDRTDQQIWDALERTFLTKAISKFKPKKLHTDVVENGDN	1277
Db	1270	LSIIPQEPVLFSGTVRSNLDPNQYTEDQIWDALERTHMKECIAQLPLKLESEVMENGDN	1329
QY	1278	FSVGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGCCTVLVIAHRVT	1337
Db	1330	FSVGERQLLCIARALLRHCKILILDEATAAMDTTETDLLIQETIREAFADCTMLTTIAHRLH	1389
QY	1338	TVLNCDHILVMNGKVVEFDRREVLRKKPGSLFAALMATATSSL	1381
Db	1390	TVLGSDRIMVLAQQQVVEFDTPSVLLSNDSSRFYAMFAAAENKV	1433

RESULT 4
MRP4_HUMAN
ID MRP4_HUMAN
AC O15439; Q9Y6J2; PRT; 1325 AA.

DT	15-JUL-1998	(Rel. 36, Created)	
DT	16-OCT-2001	(Rel. 40, Last sequence update)	
DT	15-JUN-2002	(Rel. 41, Last annotation update)	
DE	Multidrug resistance-associated protein 4 (MRP/cMOAT-related ABC transporter) (Multi-specific organic anion transporter-B) (MOAT-B).		
GN	ABCC4 OR MRP4.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=98324262; PubMed=9661885;		
RA	Lee K., Belinsky M.G., Beil D.W., Testa J.R., Kruh G.D.;		
RT	"Isolation of MOAT-B, a widely expressed multidrug resistance-associated protein/canalicular multispecific organic anion transporter-related transporter.";		
RT	Cancer Res. 58:2741-2747(1998).		
RN	[2]		
RP	SEQUENCE OF 1155-1316 FROM N.A.		
RC	TISSUE=Brain;		
RX	MEDLINE=97413640; PubMed=9270026;		
RA	Kool M., de Haas M., Scheffer G.L., Scheper R.J., van Eijk M.J.,		
RA	Juijn J.A., Baas F., Borst P.;		
RT	"Analysis of expression of cMOAT (MRP2), MRP3, MRP4, and MRP5, homologues of the multidrug resistance-associated protein gene (MRP1), in human cancer cell lines.";		
RT	Cancer Res. 57:3537-3547(1997).		
RL	-!- FUNCTION: MAY BE AN ORGANIC ANION PUMP RELEVANT TO CELLULAR DETOXIFICATION.		
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-!- TISSUE SPECIFICITY: WIDELY EXPRESSED, WITH PARTICULARLY HIGH LEVELS IN PROSTATE, BUT IS BARELY DETECTABLE IN LIVER.		
CC	-!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.		
CC	-----		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; AF071202; AAC27076.1; -.		
DR	EMBL; U83660; AAB71757.1; -.		
DR	HSSP; P13569; 1NBD.		
DR	Genew; HGNC:55; ABCC4.		
DR	MIM; 605250; -.		
DR	InterPro; IPR003593; AAA_ATPase.		
DR	InterPro; IPR003439; ABC_transportr.		
DR	InterPro; IPR001140; ABCtranprtTM.		
DR	Pfam; PF00005; ABC_tran; 2.		
DR	Pfam; PF00664; ABC_membrane; 2.		
DR	ProDom; PD000006; ABC_transportr; 2.		
DR	SMART; SM00382; AAA; 2.		
DR	PROSITE; PS00211; ABC_TRANSPORTER; 2.		
KW	ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.		
FT	TRANSMEM 93 113 POTENTIAL.		
FT	TRANSMEM 136 156 POTENTIAL.		
FT	TRANSMEM 207 227 POTENTIAL.		
FT	TRANSMEM 228 248 POTENTIAL.		
FT	TRANSMEM 328 348 POTENTIAL.		
FT	TRANSMEM 351 371 POTENTIAL.		
FT	TRANSMEM 440 460 POTENTIAL.		
FT	TRANSMEM 710 730 POTENTIAL.		
FT	TRANSMEM 771 791 POTENTIAL.		
FT	TRANSMEM 836 856 POTENTIAL.		
FT	TRANSMEM 858 878 POTENTIAL.		
FT	TRANSMEM 954 974 POTENTIAL.		
FT	TRANSMEM 977 997 POTENTIAL.		
FT	TRANSMEM 1038 1058 POTENTIAL.		
FT	NP_BIND 445 452 ATP (POTENTIAL).		
FT	NP_BIND 1075 1082 ATP (POTENTIAL).		

FT	CARBOHYD	651	651	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	690	690	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	746	746	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	754	754	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	792	792	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1176	1176	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1309	1309	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	1302	1302	H -> D (IN REF. 2).
SEQ	SEQUENCE	1325 AA;	149539 MW;	9C5750A748BB96CE CRC64;
Query Match				
Best Local Similarity 26.7%; Score 1890; DB 1; Length 1325;				
Matches 462; Conservative 251; Mismatches 513; Indels 122; Gaps 29;				
QY	83	POPLDNAGLFSYLVSWLTPMLIQSLRSRLDENTIPPLSVHSDAKNVQRLHRLWEEVS	142	
Db	10	PNPLQDANICSRVFFWNLNPLFKIGHKRRLEEDDMYSVLPEDRSQHLGEELQGFWDKEVL	69	
QY	143	R--RGIEKASVLLVMLR-FQRTRLIFDALGICFCI---ASVLGPILIPKILEYSE--E	194	
Db	70	RAENDAQKPSLTRAIIKCYWKSIVL---LGIFTLIEESAKVIOPI-FLGKIINYFENYD	124	
QY	195	QLGNVVHVGGLCPALFELSECVKSLSFSSSWIIN-----QRTAIRPRAAVSSFAFEKLI	247	
Db	125	PMDSVALNTAYATAVLTFTCLILA-----ILHLYFYHVQCAGMRRLRVAMCHMIYRKAL	179	
QY	248	QFKSVI--HITSGEAIISFFTGDNVYLFEGVGY-----GPLVLITCASLVICSISSEYFII	299	
Db	180	RLSNMAMGKTTTGQIVNLLSNDVNKFDQVTVELHFLWAGPLQAIAYTALL-----WMEI	233	
QY	300	GYTAFIAILCYLLVFPPLAVFMTMAVKAQHTSEVSDORIRVTSEVLTCIKLIKMYTWEK	359	
Db	234	GISCLAGMAVLIILLPQSCFGKLFSSLSRKTATFTDARITMNEVITGIRIIKMYAWEK	293	
QY	360	PFAKIIEDLRRKERKLLKCGLVQSLTSITLFIPTVATAVWVLIHTSLKCLKLTASMAFS	419	
Db	294	SFSNLITNLRKKKEISKILRSSCLRGMNLAFFSASKIIVFTFTTYVLLGSVITASRVFV	353	
QY	420	MLASNLRLRSV-FFVPIAVKGLTNSKSAVMRPFKF-FLQESPVFYVQTLQDPSKALVFE	477	
Db	354	AVTLYGAVRLTVTLEFFPSAIERVSEAIIVSIRRIQTFLLLDEISQRNQLPSDGKKMVHVQ	413	
QY	478	EATLSWQQTCPGIVNGALELERNGHASEGMRPRDALGP EEGNSLGP ELHKINLVVSKG	537	
Db	414	DFTAFWDK-----ASE-----TPTLQGLSFTVRPG	438	
QY	538	MMLGVCGNTGSGKSSLLSAILEEMHLLGSGVQGSLAYVPQQAIVSGNIRENILMGGA	597	
Db	439	ELLAVGVPVAGKSSLLSAVLGELAPSHGLYSVHGRIAYVSQQPWFSCTLRSNILFGKK	498	
QY	598	YDKARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKQRISLARAVYSDRQIYLL	657	
Db	499	YEKERYEKVIKACALKKDLQLEDDGLTVIGDRGTTLSGGQKARVNLARAVYQDADIYLL	558	
QY	658	DDPLSAVDAHVGKHIFEECIKKTLRGKTVVLVTHQLQYLEFCGQIILLENKICENGTHS	717	
Db	559	DDPLSAVDAEVSRLHFLFCICQILHEKITILVTHQLQYLKAAASQILILKDGKMWQKGYT	618	
QY	718	ELMQKKGKYAQLIQMKHKEATSDMLQDTAKIAEKPKVESQALA-TSLEESLNGNAVPEHQ	776	
Db	619	EFLKSGIDFGSLRKDKDNEESEQPPVPGTPTLRNRTFSESSVWSQQSSRPSRKDGAL-ESQ	677	
QY	777	LTQ-----EEEMEGSLSWRVYHHYIOAGGYMVSCIIFFVFLI-----VFLTIFSF	825	
Db	678	DTENVPVTLSEENRSEKGVGFQAYKNYFRAGAHW---IVFIFLILNTAAQVAYVLQDW	733	
QY	826	WLSYWLEQSGGTNSSRESNGTMADLGNADNPQLSFYQLVYGLNALLLTCVGCSSGIFT	885	
Db	734	WLSYW-----ANKQSMNLNVTVNGGNGVTEKLDNLWYLGISGLTVATVLFGIARSLLVF	787	
QY	886	KVTRKASTALHNKLFNKVFCPMSFFDTIPIGRLNLCFAGDLEQLDQLLPFIFSEQFLVLS	945	
Db	788	YVLVNSSQTLHNKMFESILKAPLVFFDRNPIGRIILNRFSKDIGHLDLPLTFLDFIQTL	847	

QY	946	LMVIAVLLIVSVLSPYI---LLMGALIMVICFIYYMMFKKAIGVFKRLENYSRSPLEFSHI	1002	
Db	848	LQVVGVSVAVAVIPWIAIPLVPLGIIFILRRYFLETSRDV---KRLESTRSPVFSHL	904	
QY	1003	LNSLQGLSSIHVYKGTEDFISQFKRLTDAQNNYLLLLFLSSSTRMNLRLLEIMTNLVTLAVA	1062	
Db	905	SSSLQGLWMTIRAYKAERCQELFDAHODLHSEAWFLFTTSRFAVRLDAICAMFVIIIA	964	
QY	1063	LFVAFGISSTPYSEFKV-MAVNIIVLQLASSFQATARIGLETEAQTAVERILOYMKMCVSE	1121	
Db	965	-FGSLILAKTLDAGQVGLALSALTLMGMFQWCVRQSAEVENMMISVERVIEYTDL-EKE	1022	
QY	1122	APLHMEGTSCPQGWQPHGEIIFQDYHMKYRDNTPTVLHGINLTIRGHEVVGVIGRTGSGK	1181	
Db	1023	APWEYQKRP-PPAWPHEGVIIFDNVNFMYSPPGPLVLKHLTALIKSQEKVGIVGRTGAGK	1081	
QY	1182	SSGLMALFRLVEPMAGRILIDGVDICISIGLEDLRSKLSVIPQDPVLLSGTIRFNLDPFDR	1241	
Db	1082	SSLISALFRLSEP-EGKIWDKILTTEIGLHDLRKKMSIIPQEPVLFTGTMRKNLDPFKE	1140	
QY	1242	HTDQOIWDALERTFLTKAISKFPKKLHTDVVENGNGFSGVGERQLLCIARAVLRNSKIILI	1301	
Db	1141	HTDEELWNALQEVQLKETIEDLPCKMDTELAESGSNFSVGORQLVCLARAILRKNQILII	1200	
QY	1302	DEATASIDMETDTLIQRTIREAFQGCCTVLVIAHRVTTVLNCDHILVMNGKVVEFDRPEV	1361	
Db	1201	DEATANVDPRTDELIQKKIREKFAHCTVLTIAHRLNTIIDSCKIMVLDSGRLKEYDEPYV	1260	
QY	1362	LRKKPGSLF-----AALMATA	1377	
Db	1261	LLQNKESLPHYKVMVQQLGKAEAAALTETA	1288	
RESULT 5				
MRP3_RAT				
ID	MRP3_RAT	STANDARD;	PRT;	1522 AA.
AC	O88563;	O88270;		
DT	15-JUL-1999	(Rel. 38, Created)		
DT	15-JUL-1999	(Rel. 38, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Canalicular multispecific organic anion transporter 2 (Multidrug			
DE	resistance-associated protein 3) (MRP-like protein-2) (MLP-2).			
GN	ABCC3 OR CMOAT2 OR MRP3 OR MLP2.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=99292429; PubMed=10362653;			
RA	Ortiz D.F., Li S., Iyer R., Zhang X., Novikoff P., Arias I.M.;			
RT	"MRP3, a new ATP-binding cassette protein localized to the canalicular			
RT	domain of the hepatocyte."			
RL	Am. J. Physiol. 276:G1493-G1500(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley; TISSUE=Colon;			
RX	MEDLINE=98279126; PubMed=9614210;			
RA	Hirohashi T., Suzuki H., Ito K., Ogawa K., Kume K., Shimizu T.,			
RA	Sugiyama Y.;			
RT	"Hepatic expression of multidrug resistance-associated protein-like			
RT	proteins maintained in eisa hyperbilirubinemic rats.";			
RL	Mol. Pharmacol. 53:1068-1075(1998).			
CC	-!- FUNCTION: MAY ACT AS AN INDUCIBLE TRANSPORTER IN THE BILIARY AND			
CC	INTESTINAL EXCRETION OF ORGANIC ANIONS.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-!- TISSUE SPECIFICITY: LUNG AND INTESTINE, LOW IN LIVER. HIGHER IN			
CC	LIVER OF EISAI HYPERBILIRUBINEMIC RATS AND TR(-) MUTANT RATS.			
CC	-!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			

Db 367 AVALIQSICLOTYFHCNFCNLGMCVGTVMATVYKKKALTISNLAKRQYTIGETVNLMQSDA 426
QY 269 NYLFEGVCYGPLVLITCASLVICSISSEYFI---IGYTAFIATILCYLLVFPPLAVFMRMAV 325
Db 427 OKLMDVTNFIHLVWSSVLQIV--LSIYFLWVELGPSVLAGVGMVLLIPVNGILATKNR 483
QY 326 KAQHHTSEVSDQIRIVTSEVLTICIKLIKMYTWKPFKAKTIEDLRRKERKLLKCGLVQSL 385
Db 484 NIOFKNMKYKDKRLRIMNEILSGMKILKYFAWEPSEFKDQVHNLKKELKNLRTFAYMQSV 543
QY 386 TSITLFIPTVAT---AVWLIHTSLKLKLTASMAFSMLASNLNLLRLSVFFVPPIAVKGL 441
Db 544 VMFLLYLTPVLVSVTTFSVIVLVDSN--NILDAEKAFTSITLFNILRFPMSMLPNVISAM 601
QY 442 TNSKSAVMRPFKKFF-----LOESPVYVQTLQDPSKALVFEEATLSWQQTCPGIVN 492
Db 602 LOASVSVDRLEKYLSGDDLDTSAIQRDPNF-----DKAVQFSEASFTW----- 644
QY 493 GALELERNGHASEGMRPRDALGPEEGNSLGPHELKINLVVSKGMMGLVCGNQTSGKSS 552
Db 645 -----DRN-----LEPTIRNVNLDIMPQQLVAVVGTVGSGKSS 677
QY 553 LLSAILEEMHLLGSGVGOGLAYVPPQAAWIVSGNIRENILMGGAYDKARYLQVLHCCSL 612
Db 678 LMSAMLGEMENVHGHIITIKGTAYVVPQQSWIQNGTIKNDTLFGAEFDERRYQVLEACAL 737
QY 613 NRDLLELLPFGDMTEIGERGLNLSGGQKQORISLARAVYSDRQIYLLDDPLSAVDAAHVKKHI 672
Db 738 LPDLEILPGGDLAEIGEKGINLSGGQKQORISLARASQNSDIYILDDPLSAVDAAHVKKHI 797
QY 673 FEBECI--KKTLRGKTVVLVTHQLQYLEFCGQIILLENKIKCENGTHSELMOKKGKYAQLI 730
Db 798 FNKVLGPNGLLNGKTRLLVTHSLHFLPOVDEIVVVVENGITILEKGSYSSLLAKKGVFAKNL 857
QY 731 QNMHKEATSDMLQDTAKIABKPKVESQALATSLSE----- 765
Db 858 KMFVKHTDSEGEVTVNDGSEEDDDDDSGLISSIEEPEDSISLTLKRENSLHRTLRSRR 917
QY 766 -----SLNGNAVPEHQ-----LTQEEEMEESLSWRVYHHYIOAAGGYMV 805
Db 918 SSGRRLKSLKNSLKAQNGKTPKEEVVKGQKLIKKEPFMETGKVKFSIYLYKLOATIGWCSI 977
QY 806 SCIIFFVVVLIVFLTIFSPWLSYWLQEGSGTNSRFSNGTMADLGNADNPQLSFYQLV 865
Db 978 VGIIFAYVLNSVAF-IGSNLWLSAW-----TSDSNTYNGT-----NYPASQRDLR 1021
QY 866 YGLNALLLICVGV-----CSSGIFTKVTRKASTALHNKLFNKVFCRPMSPFDTPIG 917
Db 1022 IGIFGVGLGAOGLTVLVASFWSAG-----CAHASNILHKQLLNILRAPMSFNTTPIG 1076
QY 918 RLLNCFAGDLEQLDQLPIFSEQFLVLSLMVIAVLIVSVLSP---YILLMGAIMVICF 974
Db 1077 RIVNRFAGDISTVDDTLPOSLRSWMMCFIAIISTLMICMATPVFAVIIPLAIYVAVQ 1136
QY 975 IYMMFKKAIGVFKRLENYSRSPLFESHILNSLOGLSSIHVYGKTEDFISQFKRLTDAQNN 1034
Db 1137 VFYVATSRQL---RRLDSVTRSPIYSHFTETVSGLPVIRAFEHQORFLKQNEIGIDTNQK 1193
QY 1035 YLLLFLSSTRWMALRLEIMTNLVTLAVALFVAFGISSTPYSEFKVMVAVNIVLQASSFOAT 1094
Db 1194 CVSSWITSNRWLAFLRLELVGNLVVVFSSALMMVIYRDTLSGDVVGVFLSNALNITQTLNWL 1253
QY 1095 ARIGLETEAQTAVERILOYMKMCVSEAPLHMEGTSCPQGWPQHGEIIFQDYHMKYRDNT 1154
Db 1254 VRMTSETETNIVAVERITEYIKV-ENEAP-WVTDKRPAPGWPBKGEIQFSNQVQRYRPEL 1311
QY 1155 PTVLHGINTIRGHEVGVIGRTGSGKSSGLMALFRLEVPBMAGRILIDGVDICSIGLEDL 1214
Db 1312 DLVLKGINCIDIKSMEKIGVVGRTGAGKSSLTNCLFRILEAAGGHITIDIGDIASIGLHDL 1371
QY 1215 RSKLSVIPQDPVLLSGTIRFNLDPPDRHTDQQINDALERTFLTKAISKFPKKLHTDVVEN 1274

Db 1372 RGKLTIIIPQDPVLFSGSLRMNLDPFNNYSDEEIWRALELAHLKSFVAGLQHGLSREVSEA 1431
QY 1275 GGNFSVGERQLLCCIARAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGTIVLIAH 1334
Db 1432 EDNLSIGORQLLCLGRALLRKSILVLDDEATAAVDLETDHLIQTITIRNEFSHCTVITIAH 1491
QY 1335 RVTTVLNCDHIILVMGNGKVVEFDRPEVLRKKPGSLFAALMA 1375
Db 1492 RLHTIMSDKIMVLDNGNIVEYGSPEELLESAGPF--SLMA 1530
RESULT 8
MRP2_RAT
ID MRP2_RAT STANDARD; PRT; 1541 AA.
AC Q63120; Q63145;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Canalicular multispecific organic anion transporter 1 (Multidrug
DE resistance-associated protein 2) (Canalicular multidrug resistance
DE protein).
GN ABCC2 OR CMOAT OR MRP2 OR CMRP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=96180672; PubMed=8599091;
RA Paulusma C.C., Bosma P.J., Zaman G.J.R., Bakker C.T.M., Otter M.,
RA Scheffer G.L., Scheper R.J., Borst P., Oude Elferink R.P.J.;
RT "Congenital jaundice in rats with a mutation in a multidrug
RT resistance-associated protein gene.";
RL Science 271:1126-1128(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=96279006; PubMed=8662992;
RA Buechler M., Koenig J., Brom M., Kartenbeck J., Spring H., Horie T.,
RA Keppler D.;
RT "cDNA cloning of the hepatocyte canalicular isoform of the multidrug
RT resistance protein, cMrp, reveals a novel conjugate export pump
RT deficient in hyperbilirubinemic mutant rats.";
RL J. Biol. Chem. 271:15091-15098(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RA Ito K., Suzuki H., Hirohashi T., Kume K., Shimizu T., Sugiyama Y.;
RT "Expression of the putative ATP-binding cassette region, homologous to
RT that in multidrug resistance associated protein (MRP), is hereditarily
RT defective in Eisai hyperbilirubinemic rats (EHBR).";
RL Int. Hepatol. Commun. 292:292-299(1996).
CC -!- FUNCTION: MEDIATES HEPATOBIILIARY EXCRETION OF NUMEROUS ORGANIC
CC ANIONS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: MAINLY EXPRESSED IN THE LIVER.
CC -!- DISEASE: DEFECTS IN ABCC2 ARE A CAUSE OF HEREDITARY CONJUGATED
CC HYPERBILIRUBINEMIA (EHBR).
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L49379; AAC42087.1; -.
DR EMBL: X96393; CAA65257.1; -.
DR EMBL: D86086; BAA13016.1; -.
DR HSSP: P13569; 1NBD.

DR	Pfam; PF00005; ABC_tran; 2.	
DR	Pfam; PF00664; ABC_membrane; 2.	
DR	ProDom; PD000006; ABC_transportr; 2.	
DR	SMART; SM00382; AAA; 2.	
DR	PROSITE; PS00211; ABC_TRANSPORTER; 1.	
KW	Hypothetical protein; ATP-binding; Transmembrane; Glycoprotein; Transport.	
KW		
FT	DOMAIN 1 25	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM 26 46	1 (BY SIMILARITY).
FT	DOMAIN 47 65	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM 66 85	2 (BY SIMILARITY).
FT	DOMAIN 86 90	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM 91 104	3 (BY SIMILARITY).
FT	DOMAIN 105 116	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM 117 137	4 (BY SIMILARITY).
FT	DOMAIN 138 154	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM 155 175	5 (BY SIMILARITY).
FT	DOMAIN 176 259	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM 260 280	6 (BY SIMILARITY).
FT	DOMAIN 281 310	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM 311 331	7 (BY SIMILARITY).
FT	DOMAIN 332 387	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM 388 408	8 (BY SIMILARITY).
FT	DOMAIN 409 411	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM 412 432	9 (BY SIMILARITY).
FT	DOMAIN 433 495	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM 496 516	10 (BY SIMILARITY).
FT	DOMAIN 517 539	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM 540 560	11 (BY SIMILARITY).
FT	DOMAIN 561 910	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM 911 931	12 (BY SIMILARITY).
FT	DOMAIN 932 968	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM 969 990	13 (BY SIMILARITY).
FT	DOMAIN 991 1033	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM 1034 1054	14 (BY SIMILARITY).
FT	DOMAIN 1055 1055	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM 1056 1076	15 (BY SIMILARITY).
FT	DOMAIN 1077 1147	CYTOPLASMIC (BY SIMILARITY).
FT	TRANSMEM 1148 1168	16 (BY SIMILARITY).
FT	DOMAIN 1169 1172	EXTRACELLULAR (BY SIMILARITY).
FT	TRANSMEM 1173 1193	17 (BY SIMILARITY).
FT	DOMAIN 1194 1478	CYTOPLASMIC (BY SIMILARITY).
FT	NP_BIND 631 638	ATP (POTENTIAL).
FT	NP_BIND 1273 1280	ATP (POTENTIAL).
FT	CARBOHYD 86 86	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 139 139	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE 1478 AA; 166937 MW; 6C59F43105EB7187 CRC64;	
Query Match 24.5%; Score 1734.5; DB 1; Length 1478;		
Best Local Similarity 31.9%; Pred. No. 1.6e-103;		
Matches 444; Conservative 263; Mismatches 523; Indels 161; Gaps 35;		
Qy	58 VPP---WGKYDAALRTMIPRPKPRFPAPQPLDNAGLFSYLTVSWLTPLMIQLSRLD 113	
Db	167 VPPANRVWYPDDAAELEETGLRPS-RF-----TYANIFSRISFGWLSPLMKFGYRNYLT 219	
Qy	114 EN---TIPPLSVHDASDKNVQRLHRLWEEVSRR-----GI-----EKASVLLVMLERF 158	
Db	220 ESDAWSLPPA---ERSSNLTIVFEKNWISHAKKKSSLYMMGVLFNLHWKLTVVIIVL-- 274	
Qy	159 QRTRLIFDALLGICFCIASVLGPILIIPKILEYSEEQLGNVHVGVGLCFALFLSECVKSL 218	
Db	275 ---KLQDV---VAFIQPNLIRKIVF--VSSYSSEHPQPPQVGFSLAIAMFLTNVVGTA 326	
Qy	219 SFSSSWLINQRTAIRFRAAVSSFAFEKLIQFKSVIHITSGEAISSFFTG-D-VNYL---FEG 274	
Db	327 LLQYFQLGMVLMWRWRSELIATLYRKSRLR-----LSSAARQSRVSGDIVNYSVDTQK 380	
Qy	275 VCYGPLVLITCAS---LVICSISSEYFIIGYATAFIAILCYLLVFPFLAVFMTRMAVKAQHH 330	
Db	381 VCDLTMTFLFVIVSGPFQIVLALTNLYHLVGYGALSGAFVTFLLFPCNVVVIASIFKRQNR 440	
Qy	331 TSEVSDQRIRVTSEVLTCIKLIKMYTWEKPFKIIEDLRR-KERKLEKCGLVQSLTSIT 389	

Db	441 QMKNKDARSQFMTEIINNIRSIKLYAWENIFLOKILLQLRNTRELRLMKKIGIVNTIGNFT 500	
Qy	390 LFIIPTVATA---VWVLIHTSLKLLTASMAFMSLASNLRLRSVFFVPIAVKGLTNSK 445	
Db	501 WLFAPILVSAATFGTFIVLYGKTRV-LSVDIVFACLSLFLNLQFLPTMLPIVSVSVEAS 559	
Qy	446 SAVMRFRKFF---LQESPVFYVQTLQDPS-KALVFEEATLSWQOTCPGIVNGALELERN 500	
Db	560 VAISRIYGFLTAGELDSNAVQRYPANKEPSGVCLEIKKGTFSW----- 603	
Qy	501 GHASEGMRPRDALGPEEEGNSLGPHELKINLVVSKMMLGVCGNTGSGKSSLLSAILEE 560	
Db	604 -----GPGQ--NAAEPTLRDIDFVARRGELCCIVGKVGMSKSSLLEACLG 647	
Qy	561 MHLLEGSVGVQGSLAYVPPQAWIVSGNIRENILMGGAYDKARYLQVLHCCSLNRDLELLP 620	
Db	648 MQKHSGSVPRCGSIAYAAQOPWILNATIOENILFGLGLDPEFYEKIRACCLLRDFEILA 707	
Qy	621 FGDMTETGERGLNSGGOKORISLARAVYSDRQIYLLDDPLSAVDAHVGKHFEECJ--K 678	
Db	708 DGDQTEVEGEGKISLGGQKARISLARAVYSRSDIYLLDDILSAVDQHVNRDLVRNLLGSK 767	
Qy	679 KTLRGKTVVLVTHQLOYLEFCGOIILLENGKICENGTHSELQMOKKYAQOLIQKMHKEAT 738	
Db	768 GLLRSRCVILSTNSLTVLKEASMIYMLRNCKIIESGSFTQLSSPD--SOLFQLLSEFSK 825	
Qy	739 SMLQDTAKIAEKPKEVSQALATSLEE-----SLNGNAVPEHQLTQE 780	
Db	826 KDTASSTG--ADTPLRSQSQVITSSDVTSSASRSSDTSVSNYPKATIKGTGRIRKRLTDE 883	
Qy	781 -----EEMEEGSLSWRVYHHYIQAAGGYMVSCIIFFFVVLIVFLTIFSEFWWLSYW 830	
Db	884 DNVKATGQAAAEKMERGKWKVKVWTYFKACSLFLI-FLYFLFIIGGIMNVGTNVWLKHW 942	
Qy	831 LEQSGTNSRESNGTMADLGNIDNPQLSFYQLVYGLNALL---LICVGVCSGGIFTKV 887	
Db	943 SEVNT-----QLGYNPKPYFYLGITYTLFGLLSCALISLSLTITVFCAI 986	
Qy	888 TRKASTALHNKLFNKVFRCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLM 947	
Db	987 --KSCRYLHDSMVKAVLRAPMSFFETPTGRILNRFSDDVYRDEVISRVFMFFFRNLFQ 1044	
Qy	948 VIAVLIVSVLSPYILLMGAIIMVICFIY---YMMFKKAIGVEKRLENYSRSPLFSHILN 1004	
Db	1045 IVFVLAVICYSSPMFMI---LIVPLFLYRYNQVYVYTQTSRELKRLDSVRSPLYAHFQE 1101	
Qy	1005 SLOGLSSIHVYKTEDEFISQFKRLTDAQNYYLLFLSSTRWMALRLIMTNLVTLAVALEF 1064	
Db	1102 SLGGLSTIRAYDMEDTFISENDIRVDTNHRIWFLYFSSNRWQAIRVEAIGALVVFSSAFF 1161	
Qy	1065 -VAFGISSTPYSFKV-MAVNIIVLQIASSFQATARIGLETEAQTAVERILOYMKMCVSEA 1122	
Db	1162 GVLSAVRGPNNSGLVGLSLSYAVQITQSLTFVVRQSDVDETNIIVSVERMLEYIGL-PSEA 1220	
Qy	1123 PLHMEGTSCPQGWPOHGEIIFQDYHMKYRDNTPTVLHGINLTIRGHEVVGVIGRTGSGKS 1182	
Db	1221 PSIIPDHRPPECWPSHGAIKFDHYSVRYRENPLVLNDISVNIKPQEKIGIVGRTGAGKS 1280	
Qy	1183 SLGMALFRLVEPDMAGRILIDGVDICSIGLEDRLRSKLSVTPQDPVLLSGTIRFNLDPPDRH 1242	
Db	1281 TLTALFRLIEPTSGDIQLDDINITSIGLHDLRSRLAIIPQENQAFEGTIRENLDPNANA 1340	
Qy	1243 TDQQTWDALERTFLTKAISKFPKKLHTDVVENGNGFVSGEROLLCIARAVLRNSKIILID 1302	
Db	1341 TDEEIWHALEAASLKQFIQTLDGGLYSRVTEGGANLSSGQRLMCLTRALLTPTRVLLLD 1400	
Qy	1303 EATASIDMETDTLIQRTIREAFQCGCTVLVIAHRVTTVLNCDHILVMGNKGKVVFEFDRPEVL 1362	
Db	1401 EATAAVDVETDAIVQRTIRERFNDRTILTIAHRTINTVMDSNRILVDHGKVVVEFDS TKKL 1460	
Qy	1363 RKKPGSLFAAL 1373	

QY	895	LHNLKFNKVFRCPMSEFFDTIPIGRLLNCFAGDLEQLD- - - - QLLPIFSEQFLVLSLMVIA	950	CC	or send an email to license@isb-sib.ch).
Db	967	LNLRVAKRILHTPMSYIDTTPGLRILNRFTKDTDSLDELTESTLRMTSQF- - ANIVGVC	1024	CC	
QY	951	VLLIV- - - - -SVLSPYILLMGAIIMVICFIYYMMFKKAIGVFKRLENYSRSPLESHILNS	1005	DR	PIR; S48933; S48933.
Db	1025	VMCIVYLPWFAIPAIFLL- - - - -VIFVLIADHYQSSGREI- - - - KRLEAVQRSEFVYNNLNEV	1077	DR	HSSP; P13569; INBD.
QY	1006	LOGLSSIHVYKTEDFSQFKRLTDAQN- - NYLLFLSSSTRWMALRLIMTNLVTLAVAL	1063	DR	SGD; S0001027; YHL035C.
Db	1078	LGGMDTIKAYRSQERFLAKSDFLINKMNEAGYLVVVQ- - - RWVGIFLD- - - - MVAIAFAL	1131	DR	InterPro; IPR003593; AAA_Atpase.
QY	1064	FV- - - - -AFGISSTPYSEFKVMAVNIQLASSFQATARIGLETEAQFTAVERILOYMK	1116	DR	InterPro; IPR003439; ABC_transportr.
Db	1132	IITLLCVTRAFPIISAASVG- - - VLLTYVLQLPGLLNTILRAMTQTENDMNSAERLVTYAT	1188	DR	InterPro; IPR001140; ABCtranprtrTM.
QY	1117	MCVSEAPLHMEGTSCPQGWPOHGEIIFQDYHMKYRDNTPTVLHGINLTIRGHEVVGVGR	1176	DR	Pfam; PF00005; ABC_tran; 2.
Db	1189	ELPLEASYRKPEMTTPPEPSWPMGEIIFENVDFAYRPGLPPIVLKLNLNLIKSGEKIGICGR	1248	DR	Pfam; PF00664; ABC_membrane; 2.
QY	1177	TSGKSSLCMALFRLVEPMAGRILIDGVDCISIGLEDLRSLKSVIPQDPVLLSGTIRFNL	1236	DR	ProDom; PD000006; ABC_transportr; 2.
Db	1249	TGAGKSTINSALYRLNELTAGKILIDNVDISQLGLFDLRRKLAIIPQDPVLFRTIRKNL	1308	DR	SMART; SM00382; AAA; 1.
QY	1237	DPFDRHTDOQIWDALERTFLTKAISK- - - - -FPK- - - - -KLHTD- - - - -VVENG	1276	DR	PROSITE; PS00211; ABC_TRANSPORTER; 2.
Db	1309	DPFNERTDDELWDALVR- - - - -GGAIKDDLPVKKLOKPDENGTHGRMKHFKHLDQAVEEEGS	1365	KW	Hypothetical protein; ATP-binding; Transmembrane; Glycoprotein; Transport.
QY	1277	NFSVGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGCTVLVIAHRV	1336	FT	DOMAIN 1 33 EXTRACELLULAR (BY SIMILARITY).
Db	1366	NFSLGERQLLALTRALVRQSKILILDEATSSVDYETDGGKIQTRIVEEFGDCTILCIAHRL	1425	FT	TRANSMEM 34 54 1 (BY SIMILARITY).
QY	1337	TTVLNCDHILVMGNGKVVVEFDRPEVLRKKPGSLFAAL	1373	FT	DOMAIN 55 74 CYTOPLASMIC (BY SIMILARITY).
Db	1426	KTIVNYDRILVLEKEVAEFDTPWTLFSQEDSIFRSM	1462	FT	TRANSMEM 75 95 2 (BY SIMILARITY).
RESULT 14					
YHD5_YEAST					
ID	YHD5_YEAST	STANDARD;	PRT;	1592	AA.
AC	P38735;				
DT	01-FEB-1995 (Rel. 31, Created)				
DT	01-FEB-1995 (Rel. 31, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Probable ATP-dependent permease YHL035C.				
GN	YHL035C.				
OS	Saccharomyces cerevisiae (Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.				
OX	NCBI_TaxID=4932;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=S288C / AB972;				
RX	MEDLINE=94378003; PubMed=8091229;				
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,				
RA	Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,				
RA	Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,				
RA	Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,				
RA	Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,				
RA	Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,				
RA	Vaudin M.;				
RT	"Complete nucleotide sequence of Saccharomyces cerevisiae chromosome				
RT	VIII.";				
RL	Science 265:2077-2082(1994).				
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				
CC	-!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.				
CC	- - - - -				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				

CC	or send an email to license@isb-sib.ch).				
CC	EMBL; U11583; AAB65047.1; -.				
DR	PIR; S48933; S48933.				
DR	HSSP; P13569; INBD.				
DR	SGD; S0001027; YHL035C.				
DR	InterPro; IPR003593; AAA_Atpase.				
DR	InterPro; IPR003439; ABC_transportr.				
DR	InterPro; IPR001140; ABCtranprtrTM.				
DR	Pfam; PF00005; ABC_tran; 2.				
DR	Pfam; PF00664; ABC_membrane; 2.				
DR	ProDom; PD000006; ABC_transportr; 2.				
DR	SMART; SM00382; AAA; 1.				
DR	PROSITE; PS00211; ABC_TRANSPORTER; 2.				
KW	Hypothetical protein; ATP-binding; Transmembrane; Glycoprotein; Transport.				
KW	DOMAIN 1 33 EXTRACELLULAR (BY SIMILARITY).				
FT	DOMAIN 34 54 1 (BY SIMILARITY).				
FT	DOMAIN 55 74 CYTOPLASMIC (BY SIMILARITY).				
FT	TRANSMEM 75 95 2 (BY SIMILARITY).				
FT	DOMAIN 96 100 EXTRACELLULAR (BY SIMILARITY).				
FT	TRANSMEM 101 121 3 (BY SIMILARITY).				
FT	DOMAIN 122 131 CYTOPLASMIC (BY SIMILARITY).				
FT	TRANSMEM 132 152 4 (BY SIMILARITY).				
FT	DOMAIN 153 170 EXTRACELLULAR (BY SIMILARITY).				
FT	TRANSMEM 171 191 5 (BY SIMILARITY).				
FT	DOMAIN 192 329 CYTOPLASMIC (BY SIMILARITY).				
FT	TRANSMEM 330 350 6 (BY SIMILARITY).				
FT	DOMAIN 351 379 EXTRACELLULAR (BY SIMILARITY).				
FT	TRANSMEM 380 400 7 (BY SIMILARITY).				
FT	DOMAIN 401 465 CYTOPLASMIC (BY SIMILARITY).				
FT	TRANSMEM 466 486 8 (BY SIMILARITY).				
FT	DOMAIN 487 489 EXTRACELLULAR (BY SIMILARITY).				
FT	TRANSMEM 490 510 9 (BY SIMILARITY).				
FT	DOMAIN 511 572 CYTOPLASMIC (BY SIMILARITY).				
FT	TRANSMEM 573 593 10 (BY SIMILARITY).				
FT	DOMAIN 594 614 EXTRACELLULAR (BY SIMILARITY).				
FT	TRANSMEM 615 635 11 (BY SIMILARITY).				
FT	DOMAIN 636 989 CYTOPLASMIC (BY SIMILARITY).				
FT	TRANSMEM 990 1010 12 (BY SIMILARITY).				
FT	DOMAIN 1011 1051 EXTRACELLULAR (BY SIMILARITY).				
FT	TRANSMEM 1052 1072 13 (BY SIMILARITY).				
FT	DOMAIN 1073 1115 CYTOPLASMIC (BY SIMILARITY).				
FT	TRANSMEM 1116 1136 14 (BY SIMILARITY).				
FT	DOMAIN 1137 1137 EXTRACELLULAR (BY SIMILARITY).				
FT	TRANSMEM 1138 1158 15 (BY SIMILARITY).				
FT	DOMAIN 1159 1229 CYTOPLASMIC (BY SIMILARITY).				
FT	TRANSMEM 1230 1250 16 (BY SIMILARITY).				
FT	DOMAIN 1251 1252 EXTRACELLULAR (BY SIMILARITY).				
FT	TRANSMEM 1253 1273 17 (BY SIMILARITY).				
FT	DOMAIN 1274 1592 CYTOPLASMIC (BY SIMILARITY).				
FT	NP_BIND 702 709 ATP (POTENTIAL).				
FT	NP_BIND 1357 1364 ATP (POTENTIAL).				
SQ	SEQUENCE 1592 AA; 180924 MW; 055FB0399992ACE8 CRC64;				
Query Match					22.4%; Score 1585; DB 1; Length 1592;
Best Local Similarity					29.6%; Pred. No. 7.5e-94;
Matches					432; Conservative 278; Mismatches 575; Indels 172; Gaps 35;
QY	2	TRKRRTYWVPNSSGGLVNRGIDIG--DDMVSGLIY- - - - -KTYTLQDGWPWSQQERNPEAPG	54		
Db	201	TAHTTTIWI--STWIPIRSVYIGNIDDVPSQIFYIFEVITSTLQ- - - - -PI	244		
QY	55	RAAVPPWCKYDAALRTMIPFRPKPREPAPQPLDNAGLFSYLTVSWLTPLMJQSLRSRLDE	114		
Db	245	KLTSPIKDN- - - - -SSIIYVRDDHTSPSREHI--SSILSCITWSWITNFIWEAQKNTIKL	297		
QY	115	NTIPPLSVHDASDKNVQRLHRLWEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFC	174		
Db	298	KDIWGLSMEDYSIFILKGF- - - - -TRRNKHNNLTALFESFKTYLLIGMLWVLVNS	349		
QY	175	IASVLGPILIPKILEY--SEEQLGNVVHGVGL-CFALFLSECVKLSFSSSWIINQRTA	231		

Db 350 IVNLL-PTILMKRFLIVDNPNRSSCMNLAWLYIIGMFICRLTLAICNSQGQFVSDKIC 408

QY 232 IRFRAAVSSFAFEKLIQ---FKSVIHITSGEAISPFFTGDV-----NYLFE 273

Db 409 LRIRAILIGEIIYAKGLRRRLFTSPKTSDDSDSISANLGTIINLISIDSKYSELANYLY- 467

QY 274 GVCYGPLVLITCASLVICSISSYF-IIGYTAFAIALCYLLVFPLAVFMTRMAVKAQHHTS 332

Db 468 -----VTQAVIMIIVVVGLEFLNGLVSAFAGISIIILVMEPLNELLANLLGKFQKQTL 520

QY 333 EVSDQIRIVTSEVLTCIKLIKMYTWEKPFKAIIDELRRKERKLLKCGLVQSLSITLFI 392

Db 521 KCTDQIRISKLECLQNIRIVKYFAWERNIINEIKSIRQKELRSLKKSLVWSVTSFLWFV 580

QY 393 IPTVATAVWVLIHTSLKLK-LTASMAFSMLASLNLLRLSVFFVPIAVKGLTNSKSAVMRF 451

Db 581 TPTLVTVGTFAICTFVQHEDLNAPLAFTTLSLFTLLKTPLDQLSNMLSFNOSKVSCLKRI 640

QY 452 KKFFLQESPVFYVQTLQDPSK-ALVFEEATLSWOQTCPGIVNGALELERNGHASEGMRTP 510

Db 641 SDFLRMDDTEKYNQLTISPDKNKIEFKNATLTWNE-----DSDMNAFKLCGL--- 688

QY 511 RDALGPEEENSLGPELHKINLVVSKGMLGVCNTGSGKSSLLSAILEEMHLEGSV-- 568

Db 689 -----NIKFOIGKLNLI-----GSTGSGKSALLGLLGLNELNISGSIIV 728

QY 569 -----GVQGSLAYVPQQAIVVSGNIRENILMGGAYDKARYQLQVLHCCSLNR 614

Db 729 PSLEPKHDLIPDCEGLTNSFAYCSQSAWLLNDIVKNNIIFDNFYNEDRYNKVIDACGLKR 788

QY 615 DLELLPFGDMTEIGERGLNSGGQKQRI SLARAVYSDRQIYLLDDPLSAVDAHVGKHIFE 674

Db 789 DLEILPAGDLTEIGEKGITLSGGQKQRI SLARAVYSSAKHVLLDDCLSVAVDAVWIYE 848

QY 675 ECIKKTLL-RGKTVVLVTHQLOY-LEFCGQIILLENGKICENGTHSELMQK---KGKYAQL 729

Db 849 NCITGPLMKNRTCILVTHNVSLTLRNAHFAIVLENGKVKNQGTITELQSKGLFKEKYVOL 908

QY 730 IQMKHKEATSDMLQDTAKIAEKPKVESQALATSLEESLN--GNAVPEHQLTQEEEMEES 787

Db 909 -----SSRDSINEKNANRLKAPRKNDSQKIEPVTENINFANDFVNDGQLIEEEKSNGA 962

QY 788 LSWRVYHHYIQAAGYVMVSCIIEFFVVLIVFLTFISFWLSYWLQEGSGTNSRESNGTM 847

Db 963 ISPDVYKWLKFFGGFKALTALFALYITAQILFISQSWIRHWV---NDTNVRINAPGFA 1019

QY 848 AD-----LGNIADNPQLSFYQL-VYGLNALLLCVGCSSGIFTKVT---RKASTALH 896

Db 1020 MDTPLKGMTDSSKNKHNAFYLLTVYFLIGITQAMLG---GFKTMMTFLSGMRASRKIF 1075

QY 897 NKLENKVFCRCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLLIVS 956

Db 1076 NNLLDLVLHAQIRFEDVTPVGRIMNRFSDKIEGVQDELIPYLEVTIFCLIQCASIIFLIT 1135

QY 957 VLSPYILLMGAIIMVICFIYYMMFKKAIGVFKRLENYSRSPFLSHILNSLQGLSSIHVYG 1016

Db 1136 VITPRFLTAVIVFVLYFFVGKWLTA SRELKRLDSITKSPIFHSETLVGVCTIRAFG 1195

QY 1017 KTEDFISQFKRLTDAQNNYLLFLLSST-RWMAURLLEIMTNLVTLAVALFVAFGISSTPYS 1075

Db 1196 DERRFILENMNKID-QNNRAFFYLSVTVKWFSFRVDMIGAFIVLASGSFILLNIANDSG 1254

QY 1076 FKVMAVNIVLQLASSFQATARIGLETEAQFTFAVERILOQMKMCVSEAPLHMEGTSC---P 1132

Db 1255 LAGISLTYAILFTD GALWLVRLYSTFEMNMNSVERLKEYSSIEQENYLGHDEGRILLNE 1314

QY 1133 QGWPQHGEIIFQDYHMKYRDNTPTVLHGINLTIRGHEVVGVIGRTGSGKSSLGMAFLRV 1192

Db 1315 PSWPKDGEIEIENLSRYAPNLPVPVIRNVSKVDPQSKIGIVGRTGAGKSTIITALFRLL 1374

QY 1193 EPMAGRILIDGVDICSIGLEDLRSKLSVIPQDPVLLSGTIRFNLDPPDRHTDQIWDAL 1252

Db 1375 EPITGCIKIDGQDISKIDLVTLRRSITIIQDPILFAGTIKSNVDPYDEYDEKKIFKALS 1434

QY 1253 RTFL-----TKAISKFPKKLHTDVVENGNGFNSVGERQLLCTIAKAVLRNS 1296

Db 1435 QVNLISSHEFEVLNSEERFNSTHNKF-LNLHTEIAEGGLNLSQGERQLLFTIARSLREP 1493

QY 1297 KIILIDEATASIDMETDTLIQRTIREAFQCGCTVLVIAHRVTTVLNCDHILVMGNGKVVEF 1356

Db 1494 KIILLDEATSSIDYDSHLIQGIIRSEFNKSTILTIAHRLRSVIDYDRIIVMDAGEVKEY 1553

QY 1357 DRPEVLKRKPGSLFAAL 1373

Db 1554 DRPSELLKDERGIFYSM 1570

RESULT 15

ACC8_HUMAN

ID ACC8_HUMAN STANDARD; PRT; 1580 AA.

AC Q09428; O75948; Q16583;

DT 01-NOV-1995 (Rel. 32, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Sulfonylurea receptor 1.

GN ABCC8 OR SUR1 OR SUR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RC TISSUE=Pancreatic islets;

RA Gonzalez G., Aguilar-Bryan L., Bryan J.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBBJ databases.

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Pancreas;

RA Nishimura M., Miki T., Aizawa T., Seino S.;

RL Submitted (AUG-1998) to the EMBL/GenBank/DBBJ databases.

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Foreskin, and Brain;

RA Thomas P.T., Wohllk N., Huang E., Gagel R.F., Cote G.J.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBBJ databases.

RN [4]

RP SEQUENCE OF 1186-1580 FROM N.A.

RC TISSUE=Pancreatic islets;

RX MEDLINE=95232533; PubMed=7716548;

RA Thomas P.M., Cote G.J., Wohllk N., Haddad B., Mathew P.M., Rabl W.,

RA Aguilar-Bryan L., Gagel R.F., Bryan J.;

RT "Mutations in the sulfonylurea receptor gene in familial persistent

RT hyperinsulinemic hypoglycemia of infancy.";

RL Science 268:426-429(1995).

RN [5]

RP TOPOLOGY.

RX MEDLINE=99436117; PubMed=10506167;

RA Raab-Graham K.F., Cirilo L.J., Boettcher A.A., Radeke C.M.,

RA Vandenberg C.A.;

RT "Membrane topology of the amino-terminal region of the sulfonylurea

RT receptor.";

RL J. Biol. Chem. 274:29122-29129(1999).

RN [6]

RP REVIEW ON VARIANTS.

RX MEDLINE=99268411; PubMed=10338089;

RA Meissner T., Beinbrech B., Mayatepek E.;

RT "Congenital hyperinsulinism: molecular basis of a heterogeneous

RT disease.";

RL Hum. Mutat. 13:351-361(1999).

RN [7]

RP VARIANT PHHI VAL-715.

RX MEDLINE=96354544; PubMed=8751851;

RA Thomas P.M., Wohllk N., Huang E., Kuhnle U., Rabl W., Gagel R.F.,

RA Cote G.J.;

RT "Inactivation of the first nucleotide-binding fold of the

RT sulfonylurea receptor, and familial persistent hyperinsulinemic

RT	hypoglycemia of infancy.";	CC	RECESSIVE DISORDER CHARACTERIZED BY UNREGULATED AND HIGH INSULIN
RL	Am. J. Hum. Genet. 59:510-518(1996).	CC	SECRETION. IT CAUSES NESIDIOBLASTOSIS, A DIFFUSE ABNORMALITY OF
RN	[8]	CC	THE PANCREAS IN WHICH THERE IS EXTENSIVE, OFTEN DISORGANIZED
RP	VARIANT ALA-1368.	CC	FORMATION OF NEW ISLETS. PHHI HAS A VARIABLE CLINICAL PHENOTYPE
RX	MEDLINE=96220088; PubMed=8635661;	CC	BUT USUALLY PRESENTS DURING THE NEONATAL PERIOD OR INFANCY WITH
RA	Inoue H., Ferrer J., Welling C.M., Elbein S.C., Hoffman M.,	CC	SEIZURES, COMA, AND OFTEN LARGE BIRTH WEIGHT FOR GESTATIONAL AGE.
RA	Mayorga R., Warren-Perry M., Zhang Y., Milins H., Turner R.,	CC	-!- DISEASE: DEFECTS IN ABCC8 MAY CONTRIBUTE TO AUTOSOMAL DOMINANT
RA	Province M., Bryan J., Permutt M.A., Aguilar-Bryan L.;	CC	NON-INSULIN-DEPENDENT DIABETES MELLITUS TYPE II (MODY2 OR MODY-II)
RA	"Sequence variants in the sulfonylurea receptor (SUR) gene are	CC	(ALSO KNOWN AS NIDDM), WHICH IS CHARACTERIZED BY AN AUTOSOMAL
RT	associated with NIDDM in Caucasians.";	CC	DOMINANT MODE OF INHERITANCE, ONSET DURING CHILDHOOD (USUALLY
RT	Diabetes 45:825-831(1996).	CC	BEFORE 25 YEARS OF AGE) AND A PRIMARY DEFECT IN INSULIN SECRETION.
RL	[9]	CC	-!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MRP SUBFAMILY.
RN	VARIANT PHHI PHE-1386 DEL, AND VARIANTS GLY-1359; ALA-1368 AND	CC	-----
RP	ILE-1571.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
RP	MEDLINE=97081768; PubMed=8923011;	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
RX	Nestorowicz A., Wilson B.A., Schoor K.P., Inoue H., Glaser B.,	CC	the European Bioinformatics Institute. There are no restrictions on its
RA	Landau H., Stanley C.A., Thornton P.S., Clement J.P. IV, Bryan J.,	CC	use by non-profit institutions as long as its content is in no way
RA	Aguilar-Bryan L., Permutt M.A.;	CC	modified and this statement is not removed. Usage by and for commercial
RA	"Mutations in the sulfonylurea receptor gene are associated with	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
RT	familial hyperinsulinism in Ashkenazi Jews.";	CC	or send an email to license@isb-sib.ch).
RL	Hum. Mol. Genet. 5:1813-1822(1996).	CC	-----
RN	[10]	DR	EMBL; L78243; AAB02417.1; .
RP	VARIANT PHHI ARG-1477.	DR	EMBL; L78208; AAB02417.1; JOINED.
RX	MEDLINE=96256476; PubMed=8650576;	DR	EMBL; L78209; AAB02417.1; JOINED.
RA	Nichols C.G., Shyng S.-L., Nestorowicz A., Glaser B., Clement J.P. IV,	DR	EMBL; L78210; AAB02417.1; JOINED.
RA	Gonzalez G., Aguilar-Bryan L., Permutt M.A., Bryan J.;	DR	EMBL; L78211; AAB02417.1; JOINED.
RT	"Adenosine diphosphate as an intracellular regulator of insulin	DR	EMBL; L78212; AAB02417.1; JOINED.
RT	secretion.";	DR	EMBL; L78255; AAB02417.1; JOINED.
RL	Science 272:1785-1787(1996).	DR	EMBL; L78213; AAB02417.1; JOINED.
RN	[11]	DR	EMBL; L78214; AAB02417.1; JOINED.
RP	VARIANTS GLN-274; MET-559; ASN-809; CYS-833 AND ALA-1368.	DR	EMBL; L78215; AAB02417.1; JOINED.
RX	MEDLINE=98178572; PubMed=9519757;	DR	EMBL; L78216; AAB02417.1; JOINED.
RA	Ohta Y., Tanizawa Y., Inoue H., Hosaka T., Ueda K., Matsutani A.,	DR	EMBL; L78217; AAB02417.1; JOINED.
RA	Repunte V.P., Yamada M., Kurachi Y., Bryan J., Aguilar-Bryan L.,	DR	EMBL; L78218; AAB02417.1; JOINED.
RA	Permutt M.A., Oka Y.;	DR	EMBL; L78219; AAB02417.1; JOINED.
RT	"Identification and functional analysis of sulfonylurea receptor 1	DR	EMBL; L78220; AAB02417.1; JOINED.
RT	variants in Japanese patients with NIDDM.";	DR	EMBL; L78221; AAB02417.1; JOINED.
RL	Diabetes 47:476-481(1998).	DR	EMBL; L78222; AAB02417.1; JOINED.
RN	[12]	DR	EMBL; L78223; AAB02417.1; JOINED.
RP	VARIANTS PHHI.	DR	EMBL; L78225; AAB02417.1; JOINED.
RX	MEDLINE=98282238; PubMed=9618169;	DR	EMBL; L78254; AAB02417.1; JOINED.
RA	Nestorowicz A., Glaser B., Wilson B.A., Shyng S.-L., Nichols C.G.,	DR	EMBL; L78226; AAB02417.1; JOINED.
RA	Stanley C.A., Thornton P.S., Permutt M.A.;	DR	EMBL; L78227; AAB02417.1; JOINED.
RT	"Genetic heterogeneity in familial hyperinsulinism.";	DR	EMBL; L78228; AAB02417.1; JOINED.
RL	Hum. Mol. Genet. 7:1119-1128(1998).	DR	EMBL; L78229; AAB02417.1; JOINED.
RN	[13]	DR	EMBL; L78230; AAB02417.1; JOINED.
RP	VARIANTS PHHI PRO-1351; CYS-1419 AND TRP-1492.	DR	EMBL; L78231; AAB02417.1; JOINED.
RX	MEDLINE=98443218; PubMed=9769320;	DR	EMBL; L78232; AAB02417.1; JOINED.
RA	Verkarre V., Fournet J.-C., de Lonlay P., Gross-Morand M.-S.,	DR	EMBL; L78233; AAB02417.1; JOINED.
RA	Devillers M., Rahier J., Brunelle F., Robert J.-J., Nihoul-Fekete C.,	DR	EMBL; L78234; AAB02417.1; JOINED.
RA	Saudubray J.-M., Junien C.;	DR	EMBL; L78235; AAB02417.1; JOINED.
RT	"paternal mutation of the sulfonylurea receptor (SUR1) gene and	DR	EMBL; L78236; AAB02417.1; JOINED.
RT	maternal loss of 11p15 imprinted genes lead to persistent	DR	EMBL; L78237; AAB02417.1; JOINED.
RT	hyperinsulinism in focal adenomatous hyperplasia.";	DR	EMBL; L78238; AAB02417.1; JOINED.
RL	J. Clin. Invest. 102:1286-1291(1998).	DR	EMBL; L78239; AAB02417.1; JOINED.
RN	[14]	DR	EMBL; L78240; AAB02417.1; JOINED.
RP	VARIANT PHHI ASP-186.	DR	EMBL; L78241; AAB02417.1; JOINED.
RX	MEDLINE=99265491; PubMed=10334322;	DR	EMBL; L78242; AAB02417.1; JOINED.
RA	Otonkoski T., Aemmaelae C., Huopio H., Cote G.J., Chapman J.,	DR	EMBL; L78243; AAB02418.1; .
RA	Cosgrove K., Ashfield R., Huang E., Komulainen J., Ashcroft F.M.,	DR	EMBL; L78208; AAB02418.1; JOINED.
RA	Dunne M.J., Kere J., Thomas P.M.;	DR	EMBL; L78209; AAB02418.1; JOINED.
RT	"A point mutation inactivating the sulfonylurea receptor causes the	DR	EMBL; L78210; AAB02418.1; JOINED.
RT	severe form of persistent hyperinsulinemic hypoglycemia of infancy in	DR	EMBL; L78211; AAB02418.1; JOINED.
RT	Finland.";	DR	EMBL; L78212; AAB02418.1; JOINED.
RL	Diabetes 48:408-415(1999).	DR	EMBL; L78255; AAB02418.1; JOINED.
CC	-!- FUNCTION: PUTATIVE SUBUNIT OF THE BETA-CELL ATP-SENSITIVE	DR	EMBL; L78213; AAB02418.1; JOINED.
CC	POTASSIUM CHANNEL (KATP). REGULATOR OF ATP-SENSITIVE K+ CHANNELS	DR	EMBL; L78214; AAB02418.1; JOINED.
CC	AND INSULIN RELEASE.	DR	EMBL; L78215; AAB02418.1; JOINED.
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.	DR	EMBL; L78216; AAB02418.1; JOINED.
CC	-!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are	DR	EMBL; L78217; AAB02418.1; JOINED.
CC	produced by alternative splicing.	DR	EMBL; L78218; AAB02418.1; JOINED.
CC	-!- DISEASE: DEFECTS IN ABCC8 ARE A CAUSE OF FAMILIAL PERSISTENT	DR	EMBL; L78219; AAB02418.1; JOINED.
CC	HYPERINSULINEMIC HYPOGLYCEMIA OF INFANCY (PHHI), AN AUTOSOMAL	DR	EMBL; L78220; AAB02418.1; JOINED.

Matches 425; Conservative 269; Mismatches 544; Indels 220; Gaps 38;			
QY	55	RAAVPPWGKYDAALRTMIPFRPKPRFPAPQPLDNAGLSYLTWSWLTPLMIQSLRSLDE	114
Db	201	REVKPPEDLDGLGVRFLOPF-----VNLSSKGTVMWNAFIKTAHKKPIDL	246
QY	115	NTIPPLSVHDASDKNVORLHRLWEEVSR--RGIEKASVLLVML--RFQTRRLIFDALLG	170
Db	247	RAIGKLPAMRALTNQORLCEAFDAQVRKDIOGTQGARAIQALSHAFGR-RLVLSSTFR	305
QY	171	ICFCIASVLGPILII-----PK-----ILEYSEEQLGNVHVGVGLCFALF	210
Db	306	ILADLLGFAGPLCIFGIVDHLGKENDVFPKQTQFLGVYFVSSQEFLAN---AYVLAVLLF	362
QY	211	LSECVKLSFSSSWIIINQRTAIRFRAAVSSFAFEKLIQFK----SVIHITSGEAISFFTG	266
Db	363	LALLQRTFLOASYVAIETGINLRGAIQTKIYNKIMHLSTSNLSMGEMTAGQICNLVAI	422
QY	267	DVNYL--FEGVCYGPLVLITCASLVICSISSEYFIIGYTAFAIILCYLLVFPFLAVFMRMA	324
Db	423	DTNQLMWFFFLC--PNLWAMPVQIIVGVILLYIILGVSALIGAAVIILLAPVQYFVATKL	480
QY	325	VKAQHTSEVSDQIRIVTSEVLTCIKLIKMYTWEKPPAKIIEDLRRKERKLLKCKGLVQS	384
Db	481	SQAQRSTLEYSNERLKQTNEMLRGIKLLKLYAWENIFRTRVETTRRKEMTSLRAFAI---	537
QY	385	LTSITLFI---IPTVATAVWVLIHTSL--KLKLTASMAFSMLASNLRLSLVFFVPIAVK	439
Db	538	YTSISIFMNTAIPAAVLITFVGHVSFFKEADFPSVAFASLSLFHILVTPLFLSSVVR	597
QY	440	GLTNSKSAVMRKKFFLQESPVFYVQTLQDPSKALVFEEATLSWQOTCPG-----	489
Db	598	STVKALVSVQKLSEFL-----SSAEIREEQAPHEPTPQGPASKYQAVPL	642
QY	490	-IVNGALELERNGHASE---GMTRPRDALGPEEEGNS-----LG-----PELH	528
Db	643	RVVN-----RKRPAREDCRGLTGPLQSLVPSADGDADNCCVQIMGGYFTWTPDGIPTLS	696
QY	529	KINLVVSKGMMGLVCGNTGSGKSSLLSAILEEMHLEGSV-----	568
Db	697	NITIRIPRGQLTMIVGQVCGKSSLLLAALGEMQKVSGAVFWSSLPDSEIGEDPSPERET	756
QY	569	-----GVQGLAYVPOQAMIVSGNIRENIMGGAYDKARYLQVLHCCSLNRDLELLPFG	622
Db	757	ATDLDIRKGPVAYASQKPWLLNATVEENIIFESPFNQRYKMVIEACSLQPDIDILPHG	816
QY	623	DMEIGERGLNSGGQKORISLARAVYSDROIYLLDDPLSAVDAAHVGHKHFEECIKKTLR	682
Db	817	DQTQIGERGINLSGGQORISVARALYQHANVVFLDDPPFSALDIHLSDHLMQAGILELLR	876
QY	683	G--KTVVLVTHQYLEFCGQIILLENGKICENGTHSELMOKKGYAQLI-----	730
Db	877	DDKRTVVLVTHKQLYLPHADWIIAMKDGTIQREGTLKDFORSE---CQLFEHWKTLMNRO	933
QY	731	-QKMHKCATSDMLQDTAKIAEKPKVESQALATSLSESLNGNAVPEHQLTQEEE-----	782
Db	934	DOELEKETVTE-----RKATEPPQGLSRAM-SSRDGILLQDEEEEEEAASEEDDNLSSM	987
QY	783	-MEEGSLSWRVYHHYIOAAGGYMVSCIIFFEVVLIVFLTIFSFWWLSYWLEQG-SGTNSS	840
Db	988	LHQRAEIPWRACAKYLSAGILLLSLVFSQLLKHMVLVAIDY-WLAKWTDTSALTTPAA	1046
QY	841	RESNGTMADLGNADNPQLSFYQLVYGLNALLLICVGV--CSSGIFTKVT-----RKAST	893
Db	1047	R-----NCSLSQECTLDQTVYAMVFTVLCSLGIVLC---LVTSVTVEWTGLKVAK	1093
QY	894	ALHNKLFNKVFRCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLL	953
Db	1094	RLHRSLLNRIILAPMRFETFTPLGSILNRFSSDCNTIDQHIPSTLECLSRSTLLCVSALA	1153
QY	954	IVSVLSPYILLMGALIMVICFIYYMMFKKAIGVKFRLENYSRSPLSFHILNSLQGLSSIH	1013
Db	1154	VISYVTPVFLVALLPLAIVCYFIQKYFRVASRDLQQLDDTTQLPLLSHFAETVEGLTTIR	1213

Search completed: July 21, 2003, 09:39:24
Job time : 36 secs

QY	1014	VYGKTEDFISQFKRLTDAQNNYLLFLSSSTRWMALRLIMTNLTAVALF-----	1064
Db	1214	AFRYEARFQOKLLEYTDSNNIASLFTAAARWLEVRMEYIGACVVLIAAVTSISNSLHRE	1273
QY	1065	VAFGISSTPYSFKVMVN----IVLQLASSFQATARIGLETEAQFTAVERILQYMKMCVS	1120
Db	1274	LSAGLVGLGLTYALMVSNYLNMWVRNLA-----DMELQLGAVKRIHGLLK---T	1319
QY	1121	EAPLHMEG-----TSCPQGWPHQGEIIFQDYHMKYRDNTPTVLHGINLTIIRGHEVVGVGR	1176
Db	1320	EAESY-EGLLAPSLIPKNWPDQCKIQIONLSVRYDSSLKPVCLKHVNALISPGQKIGICGR	1378
QY	1177	TSGKSSSLGMALFRLVEPVMAGRILIDGVDDICSIGLEDLRSKLSVIPQDPVLLSGTIRFNL	1236
Db	1379	TSGKSSSFLAFFRMVDTFEGHIIIDGIDIAKLPLHLRSRLSIILQDPVLFSGTIRFNL	1438
QY	1237	DPFDRHTDQIWDALERTFLTKAISKFPKKLHTDVVENGNGNFSVGERQLLCIARAVLRNS	1296
Db	1439	DPERKCSDSLWEALEFIAQLKLVVKALPGGLDAIITEGGENFSQGRQLFCLARAFVRKT	1498
QY	1297	KIILIDEATASIDMETDTLIQRTIREAFQGCTVLVIAHRVTTVLNCDHILVMGNGKVVEF	1356
Db	1499	SIFIMDEATASIDMATENILQKVVMTAFADRTVVVTIAHRVHTILSADLVIVLKRGAILEF	1558
QY	1357	DRPEVLRKKPGSLFAALM	1374
Db	1559	DKPEKLLSRKSDSVFASFV	1576

Db 1239 RALLMKSRIVLDEATASVDAGTDKLIQEVIKTVFADATVIIIAHRLDNVRNMDRIMHLK 1298
QY 1350 NGKVVEFDRPEVLRKKPGSLF 1370
Db 1299 NGKLIINFTTPOEMFKDDWSVY 1319

RESULT 4
H89715
protein F14F4.3 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001
C;Accession: H89715
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: H89715
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1404 <STO>
A;Cross-references: GB:chr_X; PIDN:CAB05730.1; PID:g3877148; GSPDB:GN00028; CESP:F14F4.3
C;Genetics:
A;Gene: F14F4.3
A;Map position: X

Query Match 26.5%; Score 1875.5; DB 2; Length 1404;
Best Local Similarity 32.1%; Pred. No. 6.5e-123;
Matches 430; Conservative 275; Mismatches 539; Indels 97; Gaps 21;

QY 62 GKYDAALRTMIPRPKPRFP--APQPLDNAGLFSYLTVSWLTPLMIQSLRSRLDENTIPP 119
Db 48 GRYSAAVQNLIPLRTTERNKNNGSGRIDDAGLFSFVTYSWFFPYLYQAARGKLDNRQVWG 107

QY 120 LSVHDASDKNVRHLRLWEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVL 179
Db 108 CSFYDSCGLMARLEVLEWEDE-KKANAKSPSLFKVIYRFISTRWFSCAVFFFLIFGFI 166

QY 180 GPILIIPKILEYSE---EQLGNVV--HGVLGCLFALFLSECVKLSFSSSSWIINQRTAIR 233
Db 167 GPTCFIRRLIAFAENPERDEQSRIVYSYGIALVAAISVVEFARVLSYGATWAVSYRTGIR 226

QY 234 FRAAVSSFAFEKLIQFKSVIHITSGEAISEFTGDNVYLFEGVCYGPLVLITCASLVICSI 293
Db 227 VRGAVLALLYNVLSKDLGCKTESDVINIFANDGQRLFDAVTFAPLVLV--GPLVLVGG 284

QY 294 SSY--FIIGYTAFIAILCYLLVFPFLAVFMTRMVAKQAQHTSEVSDORIRVTSEVLTCKL 351
Db 285 IGYLLMVIGRNSLLGILVFFVDVIOFGLGKSMVACRNLAIVKTEKRISMMAEIIKYIRI 344

QY 352 IKMYTWEKPPFAKIIEDLRRKERKLLKCGLVQSLTSITLFIIPTVA---TAVWVLIHTSL 408
Db 345 VKMNGWEQIFSAKIDQFRKEEKVQIRKSGYAQSALAIACGPVPVVAAILTFVGVLVAGN- 403

QY 409 KLKLTASMAFSLASLNLRLSVFFVPIAVKGLTNSKSAVMRFKKFFLQESPVFYVQTLO 468
Db 404 --DLLASDAFSAITVYFVLMFGIRMIPYGRYLAEAVVAMRRIOEYILLEQYAPYPTNA 461

QY 469 DPSKALVFEETATLSWQOTCPGIVNGALELERNGHASEGMRPRDALCPHEEGNSLGPPELH 528
Db 462 E-DVVLDCQCATYTYQ-----PKAAKAPVDETKE--PTEN 493

QY 529 KI-----NLVSKGMMGLVCCNTGSGKSSLLSAILLEEMHLLGSGVVGQS-L 574
Db 494 EVIVVETPVFTCSFDKLSIKRGEHIAVIGAVCGGKSAILKAISGHMFTTDDALSVDRSQT 553

QY 575 AYVPQQAIVSGNIRENILMGGAYDKARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNL 634
Db 554 VYVPQKAWIFNGTVQDNILFGDKMNSERYKAVNGCQLTEDLTTLVSGDRTEVGERGATL 613

QY 635 SGGQKQRISLARAVYSDRQIYLLDDPLSAVDAHVKGKHIFECEIKKTLRGKTVVLVTHQLQ 694
Db 614 SGGQKARVALARAVFQTKNLYLFDDEFASLDKLLHHF----- 651

QY 695 YLEFCGQIILLLENGKICENGTHSELMQKKGKYAQLIQMHK-EATSDMLQDTAKIAEKPK 753
Db 652 -----DRVLFVEGGNIVADGNHDILYENDAYKTFVDACETYQATSGATSPCGDGPQAQA 706

QY 754 VESQALATSLSEESLNGNAVPEHQLTQEEEMEESLSWRVYHHYIQAGGYMV-SCIIFFF 812
Db 707 PLDAEILRNSEDLKGDA--DKLISDEEDMGNSTIAWRIYQYIHAAGGWPIWTCLVIGF 764

QY 813 VVLIVFLTITFSFWLSYWLEQSGSGTNSRESNGT--MADLGNIADNPQLSFYQLVYGLNA 870
Db 765 IVNVV-SNIFSTYWLSRWLKKGH-DETTITNGTEFLEMKMTSLADSPVTGFYAAVYLVAL 822

QY 871 LLLICVGVCSGGIFTKVTRKASTALHNKLFNKVFRCPMSFFDTIPIGRLLNCFAGDLEQL 930
Db 823 VVLTISGLPKACVFVKVSLTAATRLHRDMFQAVIHGATSFDFSTPTGRILNRFSDKMDEI 882

QY 931 DQLLPFSEQFLVLSLMVIAVLLIVSVLSPYILLMGAIIMVICFIYMMFKKAIGVFKRL 990
Db 883 DVKLPTAEVFLQNMITCLGFLVVITSPYFLLFAIPLFVVFVFCFRAGIRNLKRS 942

QY 991 ENYSRSPLSHILNSLQGLSSIHVYKTEDFISQFKRLTDAQNNYLLFLLSSTRWMALRL 1050
Db 943 EHISRSPLYDHSVASLEGITTIHTFQOSNRFLEVLKKHLDNCNSGAIEMFQSAMRWLAVWL 1002

QY 1051 EIMTNLVTLAVALFVAFGISSTPYSFKVMANVIVLQLASSFQATARIGLETEAQTAVER 1110
Db 1003 DLLVVVMTAIVALLTVMLTGTVSPADAGMAIAFAVQMSGIFQFAVRTQTELEAKMTSVER 1062

QY 1111 ILQYMKMCVSEAPLH-MEGTSCPQGWPOHGEIIFQDYHMKYRDNTPTVLHGINLTIRGHE 1169
Db 1063 VSYADNIPEDGEWNTROGLDIESSWPANGQINFSEVNLRYRKSHPLALNDITFEIKGGE 1122

QY 1170 VVGIVGRTSGSKSLGMALFRLVEPMAGRILIDGVDCISIGLEDLSKLSVIPQDPVLLS 1229
Db 1123 KVGIIIGRTSGKSSLANLIFRLYPVTNGTIYIDGVDIRTVGLVKLRRGISAIADQPSLFS 1182

QY 1230 GTIRFNLPDRHTDQOIWDALERTFLTKAISKFPKKLHTDVVENGNFNSVGERQLLCIA 1289
Db 1183 GTVRFNLDPSLEYSDSMIWEALEKCHLKTIVQSLDKKLEADVSHGNNFNSVGERQLFCLA 1242

QY 1290 RAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGTVLVIAHRVTTVLNCDHILVMG 1349
Db 1243 RALLMKSRIVLDEATASVDAGTDKLIQEVIKTVFADATVIIIAHRLDNVRNMDRIMHLK 1302

QY 1350 NGKVVEFDRPEVLRKKPGSLF 1370
Db 1303 NGKLIINFTTPOEMFKDDWSVY 1323

RESULT 5
JE0336
canalicular multispecific organic anion transporter - human
C;Species: Homo sapiens (man)
C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 04-Jun-1999
C;Accession: JE0336
R;Uchiumi, T.; Hinoshita, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Toh, S.; Furukawa, Biochem. Biophys. Res. Commun. 252, 103-110, 1998
A;Title: Isolation of a novel human canalicular multispecific organic anion transport t.
A;Reference number: JE0336; MUID:99032812; PMID:9813153
A;Accession: JE0336
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1527 <UCH>
A;Cross-references: GB:AF083552
C;Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homolo
C;Keywords: ATP
F;1306-1499/Domain: ATP-binding cassette homology <ABC2>

Query Match		26.1%;	Score	1845.5;	DB 2;	Length	1527;
Best Local Similarity		33.4%;	Pred.	No. 9.4e-121;			
Matches	474;	Conservative	228;	Mismatches	502;	Indels	215; Gaps 35;
QY	74	FRPKPRF-----PAQPLDNAGLFSYLTVSWLTPLMIOSLRSLRDENTIPPLSVHDAS	126				
Db	190	FREKPPFFSAKNVDPNPETSAFLSRLFFWFTKMAIYGYRHPLEEKDLSLKEEDRS	249				
QY	127	DKNVQRLHRLWEEVSRGIEKAS-----VLLVMLRFQRTLIFDALL---GI	171				
Db	250	QMVVQQLLEAWRKQEKQTA RHKASAAPGKNASGEDEVLLGARPRPKPSFLKALLATFGS	309				
QY	172	CFCIASVGLPILIIIPKILEYSEEQL-----GNVVHGVGLCFALFLSEC	214				
Db	310	SFLISAC---FKLIQDLDSFINPQLLSILIRFISNPMGSPSWWGFLVAGL-----MFLCSM	361				
QY	215	VKSLSFSSSWIINQRTAIRFRAAVSSFAFEKLIQFKSVIH--ITSGEAISFTGDNVNYLF	272				
Db	362	MQSLILQHYYHYIFVTGVKFRGTGIMGVIYRKALVTNSVKRASTVGEIVNLSMSVDAQREM	421				
QY	273	EGVCYGPLVLITCASLVICSISSYFI---IGYTAFTAILCYLLVFPL---AVFMTRMAV	325				
Db	422	DLAPF--LNLWSAPLQII-LAIYFLWQNLGPSVLAGVAFMVLIIPLNGAVAVKMRAFQV	478				
QY	326	KAQHTSEVSDQIRVTSEVLTCIKLIKMYTWEKFPAKIIEDLRRKERKLLKCGLVQSL	385				
Db	479	KQM----KLKDSRIKLMSEILNGIKVLKLYAWEPSFLKQVEGIRQOGELQLLRTAAYLHTT	534				
QY	386	TSITLFIIPTVATAV--WVLIHTSLKCLKLTASMAFSMLASNLRLRLSVFFVPITAVKGLTN	443				
Db	535	TFTWMCSPLVTLITLWVYVYVDPNNVLDAAEKAFVSVSLFNILRLPLNMLPQLISNLQT	594				
QY	444	SKSAVMRFKKFFLOE--SPVFYVQTLQDPSKALVFBEATLSWQQTCPGIVNGALELERNG	501				
Db	595	ASVSLKRIQQFLSQEELDPQSVERKTIISPGYAITTHSGTFTWAQDLP-----	641				
QY	502	HASEGMTRPDALGPREEEGNSLGPELHKINLVSKGMLGVCGNTGSGKSSLSLSAILEEM	561				
Db	642	-----PTLHSLDIOVPKGALVAVVGPVPGCGKSSLSVSALLGEM	678				
QY	562	HLLEGSVGVQGLAYVPOQAWIVSGNIRENIIIMGAYDKARYLOVLHCCSLNRDLELLPF	621				
Db	679	EKLEGKVHMGKSVAYVPOQAWIQNCTLQENVLFCKALNPKRYQQTLEACALLADLEMLPG	738				
QY	622	GDMTEIGERGLNSGGQKQRI SLARAVYSDRQIYLLDDPLSAVDHAVGKHIFEECI--KK	679				
Db	739	GQOTEIGEKGINLSGGQQRVSLARAVYSDADIFLLDDPLSAVDSHVAKHIFDHVIGPEG	798				
QY	680	TLRGKTVVLVTHQLVLEFCGQIILENGKICENGTHSELMQKKGKYAQLIQKHKEATS	739				
Db	799	VLAGKTRVLVTHGISFLPQTDFIIVLADGVSEMGYPALLQRNGSFANFLCNYPADEDQ	858				
QY	740	DMLQDT-----AKIAEKPKVES-----QALATSLSESLNG	769				
Db	859	GHLEDSWTALEGAEDKEALLIEDTLSNHTDLTDNDPVTYVVQKQFMRQJLSALSSDGEGQG	918				
QY	770	NAVPEHQ-----LTQEEEMEESLSWRVYHHYIQAAAGGYMVSCIIFFF	812				
Db	919	RPVPRRHLPSEKVVTEAKADGALTQEEKAAIGTVELSVFWDYAKAVG----LCTTLAI	974				
QY	813	VVLIV---FLTIFFWLSYWLQSGGTNSSRESN-----GTMADLGNIADNPQLSFYQL	864				
Db	975	CLLYVGQSAAGANVWLSAWTNDAMA--DSRQNNTSLRLGVYAALG-----I	1020				
QY	865	VYGLNALLLICVGCSSGIFTKVTRKASTALHNKLFNKVFCRCPMSFFDTIPIGRLNLCFA	924				
Db	1021	LOGF-LVMLAAMAMAAGGI-----QAARVLHQALLHNKIRSPQSFFDTTPSGRILNCFS	1073				
QY	925	GDLEQLDQLL-PIFSEQFLVLSLM-----VIAVLLIVSVLSPYILLMGAIIIVICFIYIM	978				
Db	1074	KDIYVVDEV LAPV-----ILMLNSFFNAISTLVVIMASTP--LFTVVILPLAVLYTL	1124				
QY	979	M---FKKAIGVFKRLNYSRSPFLSHILNSLQGLSSIHVYGKTEDF--ISQFKRLTDAQN	1033				

Db	1125	VQRFYAATSRQLKRLESVSRSPYIYSHFSETVTGASVIRAYNRSRDFEIIISDTK--VDANQ	1182				
QY	1034	NYLLLFSSSTRWMALRLLEIMTNLVTLAVALFVAFGISSTPYSEFKVMVAVNIVLQLASSPQA	1093				
Db	1183	RSCYPYIISNRWLSIGVEFVGCNCVVLFAALFAVIGRSSLNPLGLVLSVSYSLQVTFALNW	1242				
QY	1094	TARIGLETEAQTAVERILOYMKMCVSEAPLHMEGSTCPQGWPOHGEIIFQDYHMKYRDN	1153				
Db	1243	MIRMSDLESNIVAVERVKEYSK-TETEAPWVVEGSRPPEGWPPRGEVEFRNYSVRYRPG	1301				
QY	1154	TPTVLHGINLTIRGHEVVGIVGRTGSGKSSSLGMALFRLVEPMAGRILIDGVDCISIGLED	1213				
Db	1302	LDLVLRDLSLHVHGKEKVGIVGRTGAGKSSMTLCLFRILEAAKGEIRIDGLNVADIGLHD	1361				
QY	1214	LRSKLSVIPQDPVLLSGTTIRFNLDPPFDRHTDQOIWDALERTFLTKAISKFPKKLHTDVVE	1273				
Db	1362	VRSQLTIIQPDPILFSGTLRMNLDPPFGSYSEEDIMWALELSHLHTFVSSQPAGLDFQCSE	1421				
QY	1274	NGGNFSVGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGCITVLVIA	1333				
Db	1422	GGENLSVGORQLVCLARALLRKSRILVLDATAAIDLETNLIQATIRTQFDTCTCTVLTIA	1481				
QY	1334	HRVTTVLNCDHILVMGKGVVEFDRPEVLVRKKKPGSLFAA	1372				
Db	1482	HRLNTIMDYTRVLVDKGVAAEFD-----SPANLIAA	1513				
RESULT 6							
S71839							
canalicular multidrug resistance protein - rat							
C;Species: Rattus norvegicus (Norway rat)							
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001							
C;Accession: S71839							
R;Buechler, M.; Koenig, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppler, J. Biol. Chem. 271, 15091-15098, 1996							
A;Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistan							
A;Reference number: S71839; MUID:96279006; PMID:8662992							
A;Accession: S71839							
A;Status: preliminary; nucleic acid sequence not shown							
A;Molecule type: mRNA							
A;Residues: 1-1541 <BUE>							
C;Cross-references: EMBL:X96393; NID:gl292881; PIDN:CAA65257.1; PID:gl617207							
C;Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homolo							
C;Keywords: ATP; glycoprotein; nucleotide binding; p-loop; transmembrane protein							
F;100-124/Domain: transmembrane #status predicted <TM01>							
F;127-151/Domain: transmembrane #status predicted <TM02>							
F;160-187/Domain: transmembrane #status predicted <TM03>							
F;305-329/Domain: transmembrane #status predicted <TM04>							
F;354-381/Domain: transmembrane #status predicted <TM05>							
F;431-451/Domain: transmembrane #status predicted <TM06>							
F;456-476/Domain: transmembrane #status predicted <TM07>							
F;536-564/Domain: transmembrane #status predicted <TM08>							
F;574-602/Domain: transmembrane #status predicted <TM09>							
F;650-833/Domain: ATP-binding cassette homology <ABC1>							
F;667-674/Region: nucleotide-binding motif A (P-loop)							
F;966-994/Domain: transmembrane #status predicted <TM10>							
F;1018-1046/Domain: transmembrane #status predicted <TM11>							
F;1104-1132/Domain: transmembrane #status predicted <TM12>							
F;1203-1228/Domain: transmembrane #status predicted <TM13>							
F;1313-1506/Domain: ATP-binding cassette homology <ABC2>							
F;1330-1337/Region: nucleotide-binding motif A (P-loop)							
F;6,1007,1010,1011/Binding site: carbohydrate (Asn) #status predicted							

Query Match 25.6%; Score 1810.5; DB 1; Length 1541;
Best Local Similarity 32.8%; Pred. No. 2.7e-118;
Matches 446; Conservative 266; Mismatches 475; Indels 171; Gaps 33;

QY	106	QSLRSRLDEN-TIPPLSVHDASDKNVQRLHRLWEEVSRRGIEKA-----	149
Db	254	QAFQRRLOKSQRKPEATLHGLNKKQSQSQDVLVEEAKKKS-EKTTKDYKPSWLKSLFK	312
QY	150	SVLLVMLRFQRTLIFDALLGICFCIASVLGPILIIIPKILEYSEEQLGNVHVGVGLC---	206

Db 766 PLSALDAHVGQVFEKCIKRELQKTRVLVTNQLHFLSQVDRIVLVHEGTVKEEGTYEEL 825

QY 719 -----LMOKKGKYAQLIQKMHKEATSDMLQDTAKIAEKPKVESQALATSLE----- 764

Db 826 SSGPLFQRLMENAGKVEEYSEE-NGEAED-----QTAEQPVANGNTNGLQMGSDDK 878

QY 765 ESLGNNAV-PEHQLTQEEEMEESGLSWRVYHHYIQAAGGYMVSCIIFFVVLIVFLTIFS 823

Db 879 KSKEGNKKGKSVLIKQEERETGVVSWRVLKRYQDALGGAWVMMLLCYVLTVEFRVTS 938

QY 824 FWWLSYWLEQSGSTNSSRESNGTMADLGNADNPQLSFYQLVYGLNALLICVGCSSGI 883

Db 939 STWLSEWTDAG-----TPKSHGPL-----FYNLIYALLSFGQVLTLTNSYW 980

QY 884 FTKVTRKASTALHNKLFNKVFRCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPFSEQFL- 942

Db 981 LIMSSLYAAKKLHNMHLHSILRAPMSFFHTNPLGRIINREFAKDLGLDRTVAVFVNMFMG 1040

QY 943 -VLSLMVIAVL-IVSVLSPYILLMGAIMVICFIYYMMFKKAIGVFKRLENYSRSPLS 1000

Db 1041 QVSQLLSTVVLIGIVSTLSLWAIMPLLVLFYGAYLYQNTAREV---KRMDSISRSPVYA 1097

QY 1001 HILNSLOGLSSHVYGKTEDFISQFK-RLTDAQNNYLLLLFSSTRNMALRLEIMTNLVTL 1059

Db 1098 QFGEALNGLSTIRAY-KAYDRMADINGRSDNNIRFTLVNMGANRWLGIRLETGLGLMIW 1156

QY 1060 AVALFVAF--GISSTPYSFKV--MAVNIVLQASSFQATARIGLETEAQFTAVERILQY 1114

Db 1157 LTASFVAMQNGRAENQOAFASTMGLLSYALNITSLLTGVLRLASLAENSLNAVERVGN 1216

QY 1115 MKMCVSEAPLHMEGTSCPOGWPOHGEIIFQDYHMKYRDNTPTVLHGINLTIRGHEVVGIV 1174

Db 1217 IEI-PPEAPPVJENNRPPPGWSSGSIKEDVVLRYRPQLPPVLVHGVSFFIHPDVKVIV 1275

QY 1175 GRTSGKSSLGNALFRLVEPMAGRILIDGVDICSIGLEDLSKLSVIPDQPVLLSGTIRF 1234

Db 1276 GRTGAGKSSLLNALFRIVEVEKGRILIDCDVGKFGGLMDLRKVLGIIPQSPVLSGTVRF 1335

QY 1235 NLDPEDRHTDQIWDALERTFLTKAISKEPKKLHTDVVENGGNFVSGERQLLCIARAVLR 1294

Db 1336 NLDPFGEHNDADLWESLERAHCLKDTIRRNPLGLDAEVSEAGENFVSQGRQLLSLRALLR 1395

QY 1295 NSKIILIDEATASIDMETDTLIQRTIREAFQGTTLVIAHRVTTVLNCDHILVMNGKV 1354

Db 1396 RSKILVDEATAAADVVRTDALIQKTIREFEKSCTMIIAHLNLTIIDCDKILVLDGRVQ 1455

QY 1355 EFDRPEVLRKKPGSLFAALM 1374

Db 1456 EFSSPENLLSNEGSSFSKMY 1475

RESULT 9

E86428

probable ABC transporter [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: E86428

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86428

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1495 <STO>

A:Cross-references: GB:AE005172; NID:g11055818; PIDN:AAG28288.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homolo

Query Match 25.4%; Score 1795; DB 2; Length 1495;

Best Local Similarity 32.9%; Pred. No. 3.2e-117;

Matches 439; Conservative 232; Mismatches 537; Indels 126; Gaps 24;

QY 84 QPLDN-----AGLFSYLTVSWLTPMLMIQSLRSRLDENTIPPLSVHDAS 126

Db 210 EPLDNVEYDALRGGEHICPERHASIFSRYFGWITPLMQLYRKPITEKDVWQDKWDQT 269

QY 127 DKNVQRLHRLWEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLGPILIIP 186

Db 270 ETLIKRFQRCWTEE-SRR--PKPWLLRALNNSLGGRFWLAGIFKICNDLSQFVGPVILSH 326

QY 187 KILEYSEEQLGNVVHGVGLCFALF---LSECVKLSFSSSWIINORTAIRFRAAVSFA 242

Db 327 LLRSMQEGDPAWV--GYVYAFIIFVGVTLGVLCEAQYFQNVW----RVGFRLRSTLVAAI 380

QY 243 FEKLIOF--KSVIHITSGEAISFFETGDVNYL-----FEGVCYGPLVLITCASLVICSIS 294

Db 381 FHKSLRLTHEARKNFASGKVTNMITTDANALQIQSQHLGLWSAPF-----RIIVSMIL 434

QY 295 SYFIIGYTAFIAILCYLLVFPLAVFMTMAVKAQHHTSEVSDQIRVITSEVLTICIKLIK 354

Db 435 LYQQLGVASLFGSLILFLLIPLQTLTIISKMRKLTKEGLQWTDKRVGITNEILSSMDTVKC 494

QY 355 YTWEKPEAKIIEDLRRKERKLLKCGLVQSLSITLFIIPTVATAVWVLIHTSLKLKLT 414

Db 495 YAWEKSFESRIQIRNEELSWFERKAQLLSAFNSFILNSIPVVTVVSVFGVFLGGDLTP 554

QY 415 SMAFSMLASLNLRLSVFFVPIAVKGLTNSKSAVMRFKFFLQESPVPYVQ-TLQDPSKA 473

Db 555 ARAFTSLSLFAVLRFPNLMLPNLLSQVNVNANVSLORIEELLSEERILAQNPPLOPQTPA 614

QY 474 LVFEEATLSWQOTCPGIVNGALELERNGHASEGMRPRDALGPPEEGNSLGPPELHKINLV 533

Db 615 ISIKNGYFSW-----DSKTTKPTLSDINLE 639

QY 534 VSKGMMLGVCGNTSGKSSLLSAILEM-HLEGSVGQVGSLAYVPQQAIVVSGNIRENI 592

Db 640 IPVGTLVAIVGGTGEKTSLSISAMLGELSHAETTSVIRGSAVYVPQVSWIFNATVRENI 699

QY 593 LMGGAYDKARYLQVLHCCSLNRDLELPPFGDMTEIGERGLNSGGOKRISLARAVYSR 652

Db 700 LFGSDFESERYWRAIDATALQHDLDLIPGRDLTEIGERGVNISGGOKRVSMARAVSNS 759

QY 653 QIYLLDDPLSAVDAAHVGHKHFEECIKTKLRGKTVLVVTHQLQYLEFCGQIILLENGKICE 712

Db 760 DVYIFDDPLSALDAHVAHQVFDSCMKDELRGKTRVLVTNQLHFLPLMDKIILVSEGMKE 819

QY 713 NGTHSELMOKKGYAQLIQKMHK-EATSDMLQDTAKIAE-KPKVESQALATSLEESLNGN 770

Db 820 EGTVELSKSGILFKKLMENAGKMDATQEVNTNDENILKLGPTVTVDVSESNLSTKQGR 879

QY 771 AVPEHQLTQEEEMEESLSWRVYHHYIQAAGGYMVSCIIFFVVLIVFLTIFSFWWLSYW 830

Db 880 R-RRSVLIKQEERETGIISWNVLMRYKEAVGGWLWVVMILLACYLATEVLRVSSSTWLSIW 938

QY 831 LEQSGGTNSSRESNGTMADLGNADNPQLSFYQLVYGLNALLICVGCSSGIFTKVTRK 890

Db 939 TDQSTSKNYS-----GFYIVVYALLGFGQAVTFTNSFWLITSSLH 980

QY 891 ASTALHNKLFNKVFRCPMSFFDTIPIGRLLNCFAGDLEQLDQ----LLPIFSEQFLVLSL 946

Db 981 AARRLHDAMLLSSILRAPMLFFHTNPTGRVINRFSKDIDGRNVANLMMFMNLWQL-L 1039

QY 947 MVIADVLLIYVSVLSPYILLMGAIMVICFIYYMMFKKAIGVFKRLENYSRSPLSHILNSL 1006

Db 1040 STFALIGTVSTISLWAIMPLLLILFYAAYLYYQSTSREV---RRLDSVTRSPIYAQFGEAL 1096

QY 1007 QGLSSIHVYKGTEDFISQFK-RLTDAQNNYLLLLFLSSTRWMALRLEIMTNLVTLAVALFV 1065

Db 1069 SQFFVNAVKTFTITVICATTWQFIFII--IPLSVFYIYYQQYVLRTSRELRLDSITRS 1126

QY 997 PLFSHILNSLOGLSSIHVYKTEDE--ISQFKRLTDAQNNYLLLF--LSSTRWMALRLLEI 1052

Db 1127 PIYSHFOETLGLGLATVRGYSQQKRFESHINQCR----IDNNSAFYPSINANRWLAYRLEL 1182

QY 1053 MTNLVTLAVALFVAFGISSTPYSPKVMVNI--VLQLASSFOATARIGLETEAQFTAVER 1110

Db 1183 IGSIIILGAATLSVRLKQGTLTAGMVGLSLSYALQITQTLNWIVRTMTVEVETNIVSVER 1242

QY 1111 ILQYMKCVSEAPLHMEGTSPOGWPQHGEIIFQDYHMKYRDNTPTVLHGINTLIRGHEV 1170

Db 1243 IKEYADL-KSEAPLIVEGHRPPKEWPSQGDIKFNYSSTRYRPBELDVLKHNINHIKPNKEK 1301

QY 1171 VGIVGRFGSGKSSGLMALFRLVPEPMAGRILIDGVDICSIGLEDRLSKLSVIPQDPVLLSG 1230

Db 1302 VGIVGRTGAGKSSLTALFRMIEASEGNIVIDNIAINEIGLYDLRHKLSIIPQDSQVFEG 1361

QY 1231 TIRENLDPFDRHTDQOIWDALERTFLTK-AISKFPKKLHDTDVVENGNGFSVGERQLLCIA 1289

Db 1362 TVRENTDIPINQYTDEAIWRALELSHLKEHVLSMSNDGLDAQLTEGGGNLSVGORQLLCIA 1421

QY 1290 RAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGCIVLVIAHRVTTVLNCDHILVMG 1349

Db 1422 RAMLVPSKILVLDEATAAVDVEDKVVQETIRTAFKDRITILTIAHRLNTIMSDRIIVLD 1481

QY 1350 NGKVVEFDRPEVLRRKKPGSLFAALMATA 1377

Db 1482 NGKVAEFDSPGQLLSDNKSLEPSLCMEA 1509

RESULT 11

T21219

hypotheical protein F21G4.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T21219; T24002

R:Mortimore, B.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19392

A:Accession: T21219

A:Status: preliminary; translated from GB/EMBL/DBRJ

A:Molecule type: DNA

A:Residues: 1-1573 <WIL>

A:Cross-references: EMBL:Z81016; PIDN:CAB02667.1; GSPDB:GN00028; CESP:F21G4.2

A:Experimental source: clone F21G4

R:Kershaw, J.

submitted to the EMBL Data Library, March 1995

A:Reference number: Z19828

A:Accession: T24002

A:Status: preliminary; translated from GB/EMBL/DBRJ

A:Molecule type: DNA

A:Residues: 1-1573 <WI2>

A:Cross-references: EMBL:Z48621; PIDN:CAA88549.1; GSPDB:GN00028; CESP:F21G4.2

A:Experimental source: clone R07B1

C:Genetics:

A:Gene: CESP:F21G4.2

A:Map position: X

A:Introns: 21/3; 122/3; 197/3; 276/1; 654/1; 911/1; 1067/1; 1472/2; 1537/2

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 24.9%; Score 1763.5; DB 2; Length 1573;

Best Local Similarity 30.1%; Pred. No. 5.6e-115;

Matches 464; Conservative 269; Mismatches 537; Indels 271; Gaps 37;

QY 5 RTYVVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGP---WSQQRNPEAPGR----- 55

Db 133 RNYGIV-TSGGL-----FISWLV---FTISAIPELLYWIIQIVNPAFAWNWIDYP 178

QY 56 -----AAVPPWKGKYDAALRTMIPFRPKPRFPAPQPPLDNAGLFS 93

Db 179 RCIAFFIWFCCAFETYLHCYADASPEGYKYLSAARN-----PSPETTSSFLN 226

QY 94 YLTVSWLTPLMIQSLRSLRDENTTPPLSVHDASDKNVQRLHRLW-----EEEVSRRGI 146

Db 227 RITMWFNLSLCSLGVKKPLEVSDLYSLNEADTSNLLVPKWKYNLWDKQSKKFEETAARRRI 286

QY 147 -EKAS-----VLLVMLRFQTRLI 164

Db 287 GSNASRTNRRRTSSNDTTPLLNDQSTDDYGSVPAGQSTQKMPSTIWTFLFMKWDVITAM 346

QY 165 FDALIG--ICFCIASVLGPILLIPKILEYSEEQNGVNVHVGCLCFALFLSECVKLSLSPSS 222

Db 347 FVKLLSDVLLFC----NP-LLLKSLIRFTEELERPMMQGVVLAFTMFFSAELSSILLSH 400

QY 223 SWIINQRTAIRFAAVSSFAFEKLIQFKSVI--HITSGEAISFFTGDNVNYLFEGVCYGPL 280

Db 401 YFYLMYRVGTRVQTCLTAAVYRKTLRLSNAARREKTVGEIVNLMIDAIDVD-RFQQIT--PQ 457

QY 281 VLTICASLVICSISSYFI---IGYTAFIAILCYLLVFPFLAVFMTRMVAKQAQHTTSEVSDQ 337

Db 458 TMOYWSNPFQIGLALFLFQQLGVSVFSGVAVMVLLFPINPVITMIIRKWQIAQMYKYKDE 517

QY 338 RIRVTSEVLTICKLIKMYTWBKPPAKIIEDLRRKERKLEKCGLVQSLTSIT-----LF 391

Db 518 RPKMVNEVLNGIKVIKLYAWBPPEQVIEDLREQELGLIKKAAFLRTFSFMLNTASPLV 577

QY 392 IIPVATAVWVLIHTSLKLKTASMAFMSLASLNLRLSVFFVPIAVKGLTNSKSAVMRF 451

Db 578 ALSTFATFIYI---DPKNVLTPEIAFVSLTLFNQLRSPMSQVAELITQTQVVVVSNNRL 633

QY 452 KKFFLQESPVFYVQTL---QDPKALVFEETLSWQQTCPGIVNGALELERNGHASEGM 507

Db 634 KEFLVSEE--LNVEAIDHRARDNNDVICKEACLSW----- 667

QY 508 TRPRDALGPREEGNSLGPDLHKINLVVSKGMMMLGVCNGTSGKSSLLSAILEEMHLLBGS 567

Db 668 -----ESAHQVPVTLTNISFSVNRGQLVTIVGRVCAKSSMLQALMGEMEKLSGS 718

QY 568 VGVOGSLAYVPOQAWIVSGNITRENILMGGAYDKARYLQVLHCCSLNRDLLELLPFGDMTEI 627

Db 719 ISMHGRLCYVPOQPWMQNNTLRQNTITFGKFDEYFYSRVLDAICALYRDLQILPLGDNTEI 778

QY 628 GERGLNLSGGQKQKORISLARAVYSDRQIYLLDDPLSAVDAHVGHKHFEECI--KKTURGKT 685

Db 779 GEKGINLSGGQKARISLARAVYQNHDYLLDDPMSAVDAHVGSQFLGSGVIGPEGMLRNKT 838

QY 686 VVLVTHQLOYLEFCGQIILLENGKICENGTHSELMQKKGKYAQLIQMKHKE----- 736

Db 839 RILVTNELSFLEKSDLIIVMNEGKIEYSGKYDDLMO-QGAFAQLLIECEKEERERREAEA 897

QY 737 -----ATSDM-----LQDTAKTAEKPK 753

Db 898 SADEDDENSEPGGIMIGGSDSDFEYDDVMA SPIIDHVLGTSHMSTVSGIINRRRI1STH 957

QY 754 VESQALAT--SLEESLNGNAVPEHQLTQEEEMEECSLSWRVYHHYIQAAAGVMVSCIIFF 811

Db 958 KORRRLLSTTKSHTHSITSASTQTRQLTGCTERVEGRVKMDTYKYFGAMG-----MSIAVL 1013

QY 812 FVVLIVFLTIFSF---WWLSYWLEQSGSGTNSRESNGTMD-----LGNADNPQLSFYQ 863

Db 1014 FVLGMTTSTIFSMGRNLWLTDW----SNDNAARSGSNTTGOPIAIRLGVYAG----- 1061

QY 864 LVYGLNALLLICVGVCS---SGIFTKVTRKASTALHNKLFNKVRCPMSPFFDPIPIGRLL 920

Db 1062 --LGFSEIILLFGLMLSLLYGGV-----SASRNLHAPLMRNLFRVPMAFYDTPFGRIL 1113

QY 921 NCFAGDLEQLDQLPIFSEQFLVLSLMVIAVLIVSVLSPYILLMGAIIIMVICFIYMMF 980

Db 1114 NRIGKDIETVDVLLPFNVQFFAQCLLQVVSTLIIMISTP---VFGIVILPLSVMYLMVM 1170

QY 981 KKAIGV---FKRLNYSRSPLSFHSILNLSLOGLSSIHVYKGTEDFISQFKRLTDAQNNYLL 1037

Db 1171 RYIATSRQLKRLESITRSPYIYSHLSEIQSATIRAYHLVDRFCKLSETKVDSHVQCRI 1230

QY 1038 LFLSSTRWMALRLIMTNLVTLAVALFVAFGISSTPYSPKVMVAVNIVLQLAASSFOATARI 1097

Db 1231 LNYVANRWLSVRLEFIGNCIVLFSALFAALTRTTTSGVIGLSVSYALNITTVLNFAVRQ 1290
QY 1098 GLETEAQFTAVERILQYMKMCVSEAPLHME-GTSCPQGWPOHGEIIFQDYHMKYRDNTPT 1156
Db 1291 ITKLETNIVSVERVKEYAE-TETEAFWKSEPGKEPPQNPWPSEGRIVMNNYSARYRPGNL 1349
QY 1157 VLHGINLTIRGHEVVGIVGRTGSGKSSLGMAFLRLVEPMPMAGRILIDGVDICSIGLEDLRS 1216
Db 1350 VVKQLNVEIKPHEKVGVGRTGAGKSSVTLSLFRIIEAAEQIIVVDGINLAEIGLHDLRS 1409
QY 1217 KLSVIPQDPVLLSGTIRFNLDPFDRHTDQOIWDALERTFTLTKAISFPKKLHTDVVENGG 1276
Db 1410 NLTIIPQDPVLFSGTLRFNLDPFNHYSDGDIWKTKLEMANLKEFAHNEQLNYIITEGGD 1469
QY 1277 NFSVGERQLLCIARAVLRNSKIILDEATASIDMETDTLIQRTIREAFQGCTVLVIAHRV 1336
Db 1470 NISVGQROLVCLARALLRKRTRVLILDEATAAVDVSTDALIQKTIREEFANATVLTIAHRL 1529
QY 1337 TTVLNCDHILVMGNGKVVEFDRPEVLRKKPGSLFAALMATA 1377
Db 1530 NTIMDYDRIIVLNDGKVGFEFDSPANLLSNRNSEFYSMAKRA 1570
RESULT 12
DVHUAR
multidrug resistance protein (cell line H69AR) - human
N:Alternate names: multidrug resistance-associated protein (MRP)
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001
C:Accession: A44231; A37495
R:Cole, S.P.C.; Bhardwaj, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.;
Science 258, 1650-1654, 1992
A:Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer
A:Reference number: A44231; MUID:93088080; PMID:1360704
A:Accession: A44231
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'MAPTRSGTGMGRGIPATPTSPSAFRTRSSCGCLVFTSGPV',50-1531 <CO1>
A:Cross-references: GB:L05628; NID:g1835658
A:Experimental source: small cell lung carcinoma cell line H69AR
A:Note: sequence extracted from NCBI backbone (NCBIP:119851); this sequence has been cor
R:Cole, S.P.C.; Deeley, R.G.
Science 260, 879, 1993
A:Title: Multidrug resistance-associated protein: sequence correction.
A:Reference number: A37495; MUID:93262415; PMID:8098549
A:Accession: A37495
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-60 <CO2>
A:Cross-references: GB:L05628; NID:g1835658
A:Note: sequence extracted from NCBI backbone (NCBIP:131929)
C:Genetics:
A:Gene: GDB:MRP
A:Cross-references: GDB:136335; OMIM:158343
A:Map position: 16p13.1-16p13.1
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C:Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transme
F:661-844/Domain: ATP-binding cassette homology <ABCL>
F:678-685/Region: nucleotide-binding motif A (P-loop)
F:788-792/Region: nucleotide-binding motif B
F:1310-1503/Domain: ATP-binding cassette homology <ABC2>
F:1327-1334/Region: nucleotide-binding motif A (P-loop)
F:1450-1454/Region: nucleotide-binding motif B
Query Match 24.9%; Score 1759; DB 1; Length 1531;
Best Local Similarity 30.3%; Pred. No. 1.1e-114;
Matches 427; Conservative 268; Mismatches 508; Indels 204; Gaps 23;
QY 81 PAPQPLDNAGLFSYLTVSWLTPMLIQSLRSRLDENTIPPLSVHDASDKNVQRLHRLWEEE 140
Db 205 PNPCPESSASFLSRITFWITGLTIVRGYRQPLEGSDLWSLNKEDTSEQVVPVLVKNWKKE 264

QY 141 VSRRGIEKASVLLVMLRFQTRLIF-----DALL-----G 170
Db 265 CAK-----TRKQPVKVYSSKDPAPQKESKVDANEVEALIVKSPOKEWNP 312
QY 171 ICFCIASVLGPILI-----IPKILEYSEEQLGNVV-----HGVLGCFALFL 211
Db 313 LFKVLYKTGPGPYFLMSFFFKAIHDLMMFSGPQILKLIKFNVDTKAPDWQGYFYTVLLFV 372
QY 212 SECVKSLSFSSSWIINQRTAIRFRAAVSSFAFEKLIQFKSVIHITS--GEAISFFTGDVN 269
Db 373 TACLQTLVLHQYFHICFVSGMRITKTAIGAVYRKALVITNSARKSSTVGEIVNLMVSDAQ 432
QY 270 YLFEGYCYGPLVITCASLVICSISYFIIGYTAFAIILCYLLVFPPLAVFMTMAVKAQH 329
Db 433 RFMDLATYINMIWSAPLQVILALYLLWLNLGSPVLGAVAVMVLMPVNAVMAAMTKTYQV 492
QY 330 HTSEVSDQIRVTSEVLTCTIKLIKMYTWEKPFKIIEDLRRKERKLLLEKCGLVQSLTSIT 389
Db 493 AHMKSNDRIKLMNEILNGIKVLKYAWELAFKDKVLAIHQEELKVLKKSAYLSAVGTFT 552
QY 390 L----FIPTVATAVVLIHTSLKLTASMAFMSLASNLLRLSVFFVPIAVKGLTNSK 445
Db 553 WVCTPFLVALCTFAVYVTIDEN--NILDAQTAFAVSLALFNILRFPNILPMVISSIVQAS 610
QY 446 SAVMRFKKFF-----LQESPVYVQTLOD--PSKALVFEETLSWQQTCPGIVNGA 494
Db 611 VSLKRLRIFLSHEELEPDSIERRPV-----KGGGTNSITVRNATFTWARSDPPTLNG- 663
QY 495 LELERNGHASEGMRPRDALGP EEGNSLGP ELHKINLVVSKGMLGVCGNTGSGKSLL 554
Db 664 -----ITFSIPEGALVAVVGQVCGGKLSLL 588
QY 555 SAILEEMHLLGSGVQGLAYVPOQAWIVSGNIRENIMLGGAYDKARYLOVLHCCSLNR 614
Db 689 SALLAEMDKVEGHVAIKGSVAYVPOQAWIQNDSLRENILFGCQLEEPYRSVIOACALLP 748
QY 615 DLELLPFGDMTEIGERGLNLSGGQKQRI SLARAVYSDRQIYLLDDPLSAVDAHVGKHIFE 674
Db 749 DLEILPSGDRTEIGEKGVNLSGGQKQKORVSLARAVYSNADIYLFDDPLSAVDAHVGKHIFE 808
QY 675 ECI--KKTLRGKTVVLVTHQLQYLEFCGQIILLENGKICENGTHSELMQKKGYAQLIQK 732
Db 809 NVIGPKGMLKNKTRILVTHSMSYLPQVDVIVMSGGKISEMSGYQELLARDGAFAEFLRT 868
QY 733 MHKEATSDMLQDTAKIA-----EKPKVESQALAT-----SLEESLNGNAV 772
Db 869 Y---ASTEQEQDAEEENGVTGVS GPGKEAKQMGMLVTDTSAGKQLORQLSSSSSYSGDIS 925
QY 773 PEHQLTOE-----EEMEEGSLSWRVYHHYIOAAGGYMVSCIIFFFVVL 815
Db 926 RHHNSTAELQKAEAKKEETWKLM EADKAQTQGVKLSVWDYMKKAIG-----LF 973
QY 816 IVFLTITFSF-----WWLSYWLEQSGGTNSSRESNGTMDLGNADNPQLSFYQL 864
Db 974 ISFLSIFLFMCNHVSALASNYWLSLWTD DDPVINGTQ EHTKVRLSVYGALGISQGIA---- 1029
QY 865 VYGLNALLLICVGVCSGGIFTKVTRKASTALHNKLFNKVFRCPMSFFDTIPIGRLLNCEFA 924
Db 1030 VFGYS-----MAVSIGGIL-----ASRCLHVDLLHLSILRSPNPFERTPSGNLVNRF 1077
QY 925 GDLEQLDQLLP IFSEQFLVLSLMVIAVLLIVSVLSPY ILLMGAIIMVICFIYMMFKKAI 984
Db 1078 KELDTVDSMIPEVIKMFMGSLFNVI GACIVILLATPIAAIIPPLGLIYFFVQRFYVASS 1137
QY 985 GVFKRLNYSRSPFLSFHILNSLQGLSSIHVYGKTEDFISQFKRLTDAQNNYLLFLSSTR 1044
Db 1138 RQLKRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANR 1197
QY 1045 WMALRLEIMTNLVT LAVALFVAFGISSTPY SFKVMAMNIVLOLASSFOATARIGLETEAQ 1104
Db 1198 WLAVRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLOQVTTYLNWLVRMSSEMETN 1257
QY 1105 FTAVERILQYMKMCVSEAPLHMEGTSCPOGWPHGEIIFQDYHMKYRDNTPTVLHGINLT 1164

Db	1258	IVAVERLKEYSE-TEKEAPWQIOETRPPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVT	1316
QY	1165	IRGHEVVGIVGRGSGKSSSGMALFRLVEPMAGRILIDGVDCISIGLEDLRSKLSVIPQD	1224
Db	1317	INGGEKVGVIGRTGAGKSSSLTGLGFRINESAEGEIIIDGINIAKIGLHDLRFKITIIPQD	1376
QY	1225	PVLLSGTIRFNLDPRHRTDQOIWDALERTFLTAKAISKFPKKLHTDVVENGGNFSVGERQ	1284
Db	1377	PVLFSGSLRMNLDPFQSYSDSEEVWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGQRQ	1436
QY	1285	LLCIARAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGCTVLVIAHRVTTVLNCDH	1344
Db	1437	LVCLARALLRKTIILVDEATAAADVLETDLDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTR	1496
QY	1345	ILVMGNKVVEFDRPEVLRKKPGSLFA	1371
Db	1497	VIVLDKGEIQEYGAPSDLLQQRGLFYS	1523
RESULT 13			
T47840			
multi resistance protein homolog - Arabidopsis thaliana			
N;Alternate names: protein T209.140			
C;Species: Arabidopsis thaliana (mouse-ear cress)			
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000			
C;Accession: T47840			
R;Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.; submitted to the Protein Sequence Database, February 2000			
A;Reference number: 224475			
A;Accession: T47840			
A;Status: preliminary			
A;molecule type: DNA			
A;Residues: 1-1490 <NYA>			
A;Cross-references: EMBL:AL138658			
A;Experimental source: cultivar Columbia; BAC clone T209			
C;Genetics:			
A;Map position: 3			
A;Introns: 9/1; 795/3; 824/3; 1036/3; 1075/3; 1174/1; 1245/3; 1347/3; 1369/1; 1449/1			
A;Note: T209.140			
C;Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology			
Query Match 24.8%; Score 1755.5; DB 2; Length 1490;			
Best Local Similarity 32.8%; Pred. No. 1.9e-114;			
Matches 446; Conservative 244; Mismatches 495; Indels 173; Gaps 31;			
QY	77	KPRFPAPQLDNAGLFSYLTVSWLTPLMIQSLRSRLDENTIPPLSVHDA---SDKNVQR	132
Db	228	KDSYSSSSPYGNATLFQRITFSWINPLFSLGKRPLEKDDVPDIDVKDSARFCSHAFAQK	287
QY	133	LHRLWEEEVRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLGPILIIIPKILEYS	192
Db	288	LKTTKEKE---GPGNAFFYNSVLRVYVWRKAINAVFVNVNASTAYIGYLINDEFVEFLS	343
QY	193	EEQLGNVVHVGVLGCFALFLSECVKSLSFSSSWIINQR-TAIRFRAAVSSFAPEK--LIQF	249
Db	344	EKQSQSLNHGYLLALGFLTAKIVETVT-QROWIFGARQLGLRLRAALISHIYQKGLVLSS	402
QY	250	KSVIHITSGEAITSFFTGDVNYLFEQVCYGPLVLITCASLVICSISSEYFI----IGYTAFI	305
Db	403	QSRQSHTSGEIINYMSVDVQORITDFIWYVNNIWM---LPIQIFSAYILQKHLGLGALA	458
QY	306	AILCYLLVFPPLAVFMTRMMAVKAQHHTSEVSDQIRVTSEVLTCIKLIKMYTWEKPFKII	365
Db	459	ALVTTLMVMACNYPLTRLQRYQSDIMNAKDRMKATSEILKNMKILKIQAWDNQFLNKV	518
QY	366	EDLRRKERKLLKCGLVQSLTSITLFIIPTVATAVWVLIHTSLKCLKLTASMAFSMLASLN	425
Db	519	KTLRKKEYDCLWKSLRLQAFTTFILWGAPSLISVVTFTVTCMLMGVKLTAGAVLSALATFQ	578
QY	426	LLRLSVFFVPVIAVKGLTNSKSAMVRKFKFFLQESPVEYVQTLQDPDSKALVFEEATLSWQQ	485
Db	579	MLQSPIFGLPDLLSALVQSKVSADRIAS-YLQQS-----ETQKDAVEYCSKDHTELSVE-	631

QY	486	TCPGIVNGALELERNGHASEGMRPRDALGPDEEGNSLGPELHKINLVVSKGMMLGVCGN	545
Db	632	----IENGAFSWE-----SSRPTLDDIELKVKSGMKVAVCGA	666
QY	546	TSGSKSSLLSAILEEMHLLLEGSVGVQGS LAVVPPQAAWIVSGNIRENILMGAYDKARYLQ	605
Db	667	VSGKSSLLSSILGEIQKLGKTVRVSGKQAVVPQSPWILSGTIRDNILFGSMYSEKYER	726
QY	606	VLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKQRI SARAVYSDRQIYLLDDPLSAVD	665
Db	727	TVKACALIKDFELFSNGDLTEIGERGINMSGQKQRIARAVYQNAIDYLLDDPFSAVD	786
QY	666	AHVGKHIFEECIKKTLRGKTVVLVTHQLQYLEFCGQIILLENGKICENGTHSELMQKKGK	725
Db	787	AHTGRELFEDCLMGILKDKTVLYVTHQVEFLPAADLILVMQGRVMOAGKFEELLKQNI	846
QY	726	YAQLIQMKHKEATSDML-----QDTAKIAEKPKVESQALATSLEESLN---	768
Db	847	FEVLV-GAHNEALDSILSIEKSSRNFKESKDDDTASIAE-----SLOTHCDSEHNIST	898
QY	769	GNAVPEHQLTQEEEMEESLSWRVYHHYIQAA-GGYMVSCIIFFFVVL-----IVFLTIFS	823
Db	899	ENKKKEAKLVQDEETEKGVIGEVYLA YLTVKGGLLVP-----FIILAQSCFQMLQIAS	953
QY	824	FWLSYWLEQSGTNSRESNGTMADLGNADNPQSFYQLVYGLNA-----LLL	873
Db	954	NYWMA-WTAPPTAESIPKLGMRIL-----LVYALLAAGSSLCVLAARTIL	997
QY	874	ICVGVCSGSGIFTKVTTRKASTALHNKLFNKVRCPMSPFDTIPIGRLLNCFAGDLEQDQL	933
Db	998	VAIG---GLSTAET-----FFSRMLCSIFRAPMSFFDSTPTGRILNRASTDQSVLD--	1045
QY	934	LPIFSEQFLVLSLMVIAVLLIVSVLSPYILLGAIIMVICFIYYMMFKKAIGVFKRLENY	993
Db	1046	---LEMAVKLGWCAFSIIQIV-----GTIFVMSQVAWQRYTPTARELSRMSGV	1091
QY	994	SRSPLFSHILNSLOGLSSIHVYGKTEDEFISQFKRLTDAONNYLLLFSSTRWMALREIM	1053
Db	1092	ERAPILHHFAESLAGATTIRAFDQDRDRFISSNLVLIDSHSRPWFHVASAMENLSFRLLN	1151
QY	1054	TN-----LVTL-----AVALFVAFGISSTPYSFVKVMVAVNIVLQALASSFOATAR	1096
Db	1152	SHVFVAFSLVLLVTLPEGVINPSIAGLVTYGL-----SLNVLQATVIWNICNA-----	1200
QY	1097	IGLETEAQFTAVERILQYMKMCVSEAPLHNEGTCSPQGWPHQGEIIFQDYHMKYRDNTPT	1156
Db	1201	-----ENKMISVERILQYSKI-PSEAPLVIDGHRPLDNWPNVGSIVFRDLQVRYAEHFA	1254
QY	1157	VLHGINLTIRGHEVVGIVGRTSGSKSSLGMALFRLVEPMAGRILIDGVDCISIGLEDLRS	1216
Db	1255	VLKNITCEFPFGKKIGVWVRTSGSKSTLQALFRIVEPSQGTIVIDNVNIDITKIGLHDLRS	1314
QY	1217	KLSVIPQDPVLLSGTIRFNLDPPDRHTDQOIWDALERTFLTAKAISKFPKKLHTDVVENG	1276
Db	1315	RLGIIPQDPALPFDGTIRNLDP LAQYTDHETWEAIDKCQLGDVIRAKDERLDATVVENGE	1374
QY	1277	NFSVGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQCTVLVIAHRV	1336
Db	1375	NWSVGQRQLVCLGRVLLKKSNIILVDEATASVDSATDGVIOKIINQEFKDRVTVVVIAHRI	1434
QY	1337	TTVLNCDHILVMGNKVVEFDRPEVLRKKPGSLFAALM	1374
Db	1435	HTVIESDLVLVLSDGRIAEFDSPAKLLQREDSEFFSKLI	1472

RESULT 14

S71841

multidrug resistance protein, canalicular - human

C;Species: Homo sapiens (man)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001

C;Accession: S71841; S71840

R;Koenig, J.; Keppler, D.

submitted to the EMBL Data Library, August 1996
A;Reference number: S71841
A;Accession: S71841
A;Molecule type: mRNA
A;Residues: 1-1545 <KOE>
A;Cross-references: EMBL:X96395; NID:g1507819; PIDN:CAA65259.1; PID:g1514568
R;Buechler, M.; Koenig, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppler, D.
J. Biol. Chem. 271, 15091-15098, 1996
A;Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance
A;Reference number: S71839; MUID:96279006; PMID:8662992
A;Accession: S71840
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1415-1429, 'VP', 1432-1455, 'E', 1457-1545 <BUE>
A;Cross-references: EMBL:X96395
C;Genetics:
A;Gene: GDB:ABCC2; CMOAT; ABC; MRP2; CMRP; DJS
A;Cross-references: GDB:6089489; OMIM:601107
A;Map position: 10q24-10q24
C;Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C;Keywords: ATP; glycoprotein; nucleotide binding; p-loop; transmembrane protein
F;654-837/Domain: ATP-binding cassette homology <ABC1>
F;671-678/Region: nucleotide-binding motif A {p-loop}
F;1317-1510/Domain: ATP-binding cassette homology <ABC2>
F;1334-1341/Region: nucleotide-binding motif A (p-loop)

Query Match 24.8%; Score 1754; DB 1; Length 1545;
Best Local Similarity 32.3%; Pred. No. 2.5e-114;
Matches 424; Conservative 263; Mismatches 462; Indels 163; Gaps 30;

QY 138 EEEVRRRGIEKA---SVLLVMLRFQTRLIFDALLGICFCIASVLGPILIIPKILEYSEE 194
DB 302 KKDVPKSLWKALFKTEYMYLLKSFLKLKLVND-----IFTEVSPQL-LKLLISFASD 352

QY 195 QLGNNVHGVGLCFALFLSECVKSLSFSSSWIINQRTAIRFRAAVSSFAFEKLIQFKSVI- 253
DB 353 RDTYLWIGYLCAILLFTTAALIQSFCLQCYFQLCFKLGKVKVRTAIMASVYKKALTLSNLR 412

QY 254 -HITSGEASIFFTGDNVLPFEGVCYGPLVLITCASLVICSISSYFI---IGYTAFIAILC 309
DB 413 KEYTVGETVNLMSVDAQKLMDDVTNFMHMLWSSVLQIV---LSIFFLWRELGPVSLAGVG 469

QY 310 YLLVFPLAVFMTMAVKAQHHTSEVSDQRIRVTSEVLTCIKLIKMYTWKPPAKIIEDLR 369
DB 470 MVLVIPINAILSTKSTIQVKNMKNKDKRLKIMNEILSGIKILKYFAWEPSFRDQVQNLR 529

QY 370 RKERKLLKCGLVQSLT----SITLFIPTVATAVVMVLIHTSLKCLKLTASMAPSMLASLN 425
DB 530 KKELKNLLAFSQLQCVLIFVFQLTPVLVSVVTFVSFVVLVDSN--NILDQAQKAPTSITLFN 587

QY 426 LLRLSVFFVPIAVKGLTNSKSAVMRFKKF---LQESPVYVQTLQDPSKALVFEEATL 481
DB 588 ILRFPLSMLPMMISSMLQASVSTERLEKYLGGDDLDTSAIRHSCNF--DKAMQFSEASF 644

QY 482 SWQQTCPGIVNGALELERNGHASEGMRPRDALGPBEEGNSLGPHELHKINLVVSKGMMLG 541
DB 645 TWE-----HDSEATVR-----DVNLDIMAGQLVA 668

QY 542 VCGNTGSGKSSLLSAILEMHLLLEGVGVQGS LAVPQQAWIVSGNIRENIMGCGAYDKA 601
DB 669 VIGPVGSGKSSLISAMLGEMENVHGHIITKGTAYVPQQSWIQNGTIKDNILFGTEFNEK 728

QY 602 RYLOVLHCCSLNRDLELPPFGDMTEIGERGLNLSGQKQKRISLARAVYSDRQIYLLDDPL 661
DB 729 RYQQVLEACALLPDLMLPGGDLAEIGEKGINLSGQKQKRISLARATYQNLDIYLLDDPL 788

QY 662 SAVDAHVGKHFEECI--KKTLRGKTVVLVTHQLQVLEFCGQIILLENGKICENGTHSEL 719
DB 789 SAVDAHVGKHFINKVLGPNGLLKGKTRLLVTHSMHFLPQVDEIVVLNGTIVEKGSYSAL 848

QY 720 MQKKGKYAQLIQKM-----HKEAT-----SDMLQDTAKIAEKPKVESQALA----- 760
DB 849 LAKKGEFAKNLKTFLRHTGPREEATVHDGSEEEADDYGLISSVEEIPEDAASITMRRENS 908

QY 761 -----TSLEESL---NGNAVPE-----HQLTOEEEEEGSLSNRVYH 794
DB FRRTLRSRRSNGRHLKSLRNSLKRNVNSLKEDEELVKGQKLIKKEFIETGKVKFSIYL 968

QY 795 HYIQAAGGYMVSCIIFPFVVLIVFLTIFSFWMLSYVLEQSGGTNSSRESNGTMADLGNIA 854
DB EYLOAIGLFSIFFIILAFVMNSVAF-IGSNLWLSAW-----TSDSKIFNST----- 1013

QY 855 DNP--QLSFYQLVYGLNALLLICVGVCSGGIFTKVYR-----KASTALHNKILFNKVF 904
DB 1014 DYPASQRDMRVGVYGALGL-----AQGIFVIAHFWSAFGVHASNILHKQLLNNIL 1065

QY 905 RCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLIVSVLSP---Y 961
DB 1066 RAPMRFDDTPTGRIVNRFAGDIISTVDDTLPQSLRTWITCFLGIISTLVMICMATPVFTI 1125

QY 962 ILLMGAIIMVICFIYYMMFKKAIGVFKRLENYSRSLFSHILNSLQGLSSIHYGKTEDF 1021
DB 1126 IVIPLGIYVSVMFYVSTSRQL--RRLDSVTRSPIYSHFSETVSGLPVIRAFEHQORF 1182

QY 1022 ISQFKRLTDAQNNYLLLFSSTRWMALRLEIMTNLVTLAVALFVAF--GISSTPYSFKV 1078
DB 1183 LKHNEVRIDTNQKCVFSWITSNRWLAIIRLELVGNLTVFFSALMMVIYRDTLSGDTVGF-- 1240

QY 1079 MAVNIVLQLASSFOATARIGLETEAQFTAVERILQYMKMCVSEAPLHMEGTSCTPQGWPOH 1138
DB 1241 -VLSNALNITQTLNWLVRMTSEIETNIVAVERITEYTKV-ENEAP-WVTDKRPPDPWPSK 1297

QY 1139 GEIIFQDHYMKYRDNTPTVLHGINLTIRGHEVVGIVGRGTSGKSSLGMAFLRLVEPMAGR 1198
DB 1298 GKIQFNYYQVRYRPELDLVRGITCDIGSMEKIGVVGRGTGAGKSSLTNCLFRILEAAGGQ 1357

QY 1199 ILIDGVDICSIGLEDRSKLSVIPQDPVLLSGTIRFNLDPPFDRHTDQQLWDALERTFLTK 1258
DB 1358 LIIDGVDIASIGLHDLREKLTIIPODPILFSGSLRMLNLPFNYSDEEIWKALELAHLKS 1417

QY 1259 AISKFPKKLHTDVVENGNGFSGVERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQR 1318
DB 1418 FVASLQLGLSHEGTEAGNLSIGORQLCLGRALLRKSKILVLDEATAAVDLETDNLJOT 1477

QY 1319 TIREAFQGCCTVLVIAHRVTTVLCNDHILVMGNGKVVEFDRPEVLRRKKPGSLF 1370
DB 1478 TIQNEFAHCTVITIAHRLHTIMDSDKVMVLNDGKILIECGSPEELIQIPGPFY 1529

RESULT 15
E89447
protein F57C12.4 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001
C;Accession: E89447
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_-
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A;Accession: E89447
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1494 <STO>
A;Cross-references: GB:chr_X; PIDN:AAA83299.1; PID:g1118071; GSPDB:GN000028; CESP:F57C
C;Genetics:
A;Gene: F57C12.4
A;Map position: X
C;Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homolo

Query Match 24.8%; Score 1751; DB 2; Length 1494;
Best Local Similarity 31.3%; pred. No. 3.9e-114;
Matches 439; Conservative 270; Mismatches 512; Indels 180; Gaps 31;

QY 85 PLDNAGLFSYLTVSNLTPLMIQSLRSRLDENTIPPLSVHDASDKNVQRLRWEEV\$-- 142

[illegible]

Search completed: July 21, 2003, 09:42:21
Job time : 59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 21, 2003, 09:28:55 ; Search time 49 Seconds
(without alignments)
3758.211 Million cell updates/sec

Title: US-10-087-782A-31
Perfect score: 7071
Sequence: 1 MTRKRTYVVPNSSGGLVNRG.....RKKPGSLFAALMATATSSLR 1382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	7034.5	99.5	1379	22 AAB62555	Novel human transp
2	6807.5	96.3	1347	23 ABG61895	Prostate cancer-as
3	6791	96.0	1360	23 AAM51155	Human ATP-binding
4	6723.5	95.1	1331	23 AAE22722	Human adenosine tr
5	6096	86.2	1216	22 AAB62558	Novel human transp
6	5414	76.6	1063	22 AAB62549	Novel human transp
7	5218.5	73.8	1021	22 AAB62554	Novel human transp
8	4985.5	70.5	975	22 AAB62553	Novel human transp
9	4475.5	63.3	900	22 AAB62552	Novel human transp
10	4280	60.5	858	22 AAB62557	Novel human transp

11	4047	57.2	812	22	AAB62556	Novel human transp
12	3598	50.9	705	22	AAB62548	Novel human transp
13	3365	47.6	659	22	AAB62547	Novel human transp
14	3343	47.3	1363	22	AAE08075	Human transporter-
15	3208.5	45.4	1332	22	AAE08078	Human transporter-
16	2944	41.6	606	23	AAE21158	Human TRICH-2 prot
17	2803	39.6	1216	22	AAE08076	Human transporter-
18	2786	39.4	1247	22	AAE08077	Human transporter-
19	2665.5	37.7	1453	22	AAB47021	Multidrug-resistan
20	2663.5	37.7	1437	19	AAW80597	Human multidrug re
21	2663.5	37.7	1437	21	AAH10225	Human MRP-beta pro
22	2661.5	37.6	1437	20	AAI43542	A human MPR-relate
23	2659.5	37.6	542	22	AAB62551	Novel human transp
24	2426.5	34.3	496	22	AAB62550	Novel human transp
25	2129	30.1	1182	22	ABC25897	Novel human diagno
26	1894	26.8	1325	23	ABG61820	Prostate cancer-as
27	1886	26.7	1325	20	AAI43541	A human MPR-relate
28	1848.5	26.1	1527	23	AAU91309	Human protein NOV1
29	1844.5	26.1	1527	20	AAI43543	A human MPR-relate
30	1844	26.1	1261	22	AAU69824	Human prostate CDN
31	1844	26.1	1261	22	AAM01179	Human prostate-spe
32	1844	26.1	1261	22	AAG99064	Human prostate-spe
33	1844	26.1	1261	23	ABH95284	Human P510S protei
34	1842	26.1	1528	18	AAW33363	Human multidrug re
35	1810.5	25.6	1541	18	AAW33361	Rat canalicular mu
36	1795	25.4	1495	23	ABB91189	Herbicidally activ
37	1794.5	25.4	1528	19	AAW57487	Murine multidrug r
38	1794.5	25.4	1528	20	AAW99895	Mouse multidrug re
39	1794.5	25.4	1528	20	AAW74472	Mouse multidrug re
40	1794.5	25.4	1528	21	AAI78874	Murine multidrug r
41	1794.5	25.4	1528	21	AAI55800	Murine multidrug r
42	1781.5	25.2	1621	19	AAW60162	Arabidopsis thalia
43	1780.5	25.2	1622	19	AAW60163	Arabidopsis thalia
44	1775	25.1	1545	18	AAW33362	Human canalicular
45	1767	25.0	1531	19	AAW57486	Human MRP variant

ALIGNMENTS

RESULT 1

AAB62555

ID AAB62555 standard; Protein; 1379 AA.

XX

AC AAB62555;

XX

DT 23-JUL-2001 (first entry)

XX

DE Novel human transporter protein (NHP).

XX

KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;

KW gene therapy; screening.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Misc-difference 86..87

FT /note= "the corresponding DNA encodes 3 amino acid

FT residues in between the above positions which

FT is not indicated in the present sequence"

XX

PN WO200132706-A2.

XX

PD 10-MAY-2001.

XX

PF 31-OCT-2000; 2000WO-US29852.

XX

PR 02-NOV-1999; 99US-0163018.

XX

PA (LEXI-) LEXICON GENETICS INC.

XX

PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;

PI Sands AT;

ID XX AAM51155 standard; Protein; 1360 AA.
AC XX AAM51155;
DT 10-JUN-2002 (first entry)
XX
DE Human ATP-binding cassette transporter 44589.
XX
KW ATP-binding cassette transporter; ABC transporter 44589; human;
KW cancer; osteopathic; immunomodulatory; cardiovascular; analgesic;
KW antidiabetic; cytostatic; antileukaemic; hepatic; vaccine;
KW diagnosis; therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 515..686
FT /note= "first ABC transporter ATP cassette domain"
FT Domain 1146..1329
FT /note= "second ABC transporter ATP cassette domain"
FT Region 1..162
FT /note= "cytoplasmic region"
FT Region 163..445
FT /note= "first ABC transporter transmembrane region"
FT Region 163..185
FT /note= "transmembrane helix"
FT Region 186..198
FT /note= "extracellular region"
FT Region 199..215
FT /note= "transmembrane helix"
FT Region 216..282
FT /note= "cytoplasmic region"
FT Region 283..303
FT /note= "transmembrane helix"
FT Region 304..309
FT /note= "extracellular region"
FT Region 310..333
FT /note= "transmembrane helix"
FT Region 334..352
FT /note= "cytoplasmic region"
FT Region 353..369
FT /note= "transmembrane helix"
FT Region 370..395
FT /note= "extracellular region"
FT Region 396..416
FT /note= "transmembrane helix"
FT Region 417..780
FT /note= "cytoplasmic region"
FT Region 784..1073
FT /note= "second ABC transporter transmembrane region"
FT Region 781..805
FT /note= "transmembrane helix"
FT Region 806..841
FT /note= "extracellular region"
FT Region 842..863
FT /note= "transmembrane helix"
FT Region 864..918
FT /note= "cytoplasmic region"
FT Region 919..935
FT /note= "transmembrane helix"
FT Region 936..941
FT /note= "extracellular region"
FT Region 942..958
FT /note= "transmembrane helix"
FT Region 959..1029
FT /note= "cytoplasmic region"
FT Region 1030..1047
FT /note= "transmembrane helix"
FT Region 1048..1051
FT /note= "extracellular region"
FT Region 1052..1069
FT /note= "transmembrane helix"
FT Region 1070..1360

FT Binding-site /note= "cytoplasmic region"
FT 522..529
FT /note= "ATP/GTP binding site motif A (P-loop)"
FT 1153..1160
FT /note= "ATP/GTP binding site motif A (P-loop)"
FT 616..626
FT /note= "ABC transporter family signature"
FT 1256..1270
FT /note= "ABC transporter family signature"
FT 11..14
FT /note= "Asn is N-glycosylated"
FT 611..614
FT /note= "Asn is N-glycosylated"
FT 691..694
FT /note= "Asn is N-glycosylated"
FT 816..819
FT /note= "Asn is N-glycosylated"
FT 822..825
FT /note= "Asn is N-glycosylated"
FT 970..973
FT /note= "Asn is N-glycosylated"
FT 1140..1143
FT /note= "Asn is N-glycosylated"
FT 1255..1258
FT /note= "Asn is N-glycosylated"
FT 257..260
FT /note= "glycosaminoglycan attachment site"
FT 3..6
FT /note= "O-phosphorylated by CAMP/CGMP protein kinase"
FT 867..870
FT /note= "O-phosphorylated by CAMP/CGMP protein kinase"
FT 1004..1007
FT /note= "O-phosphorylated by CAMP/CGMP protein kinase"
FT 2..4
FT /note= "O-phosphorylated by protein kinase C"
FT 107..109
FT /note= "O-phosphorylated by protein kinase C"
FT 126..128
FT /note= "O-phosphorylated by protein kinase C"
FT 142..144
FT /note= "O-phosphorylated by protein kinase C"
FT 526..528
FT /note= "O-phosphorylated by protein kinase C"
FT 628..630
FT /note= "O-phosphorylated by protein kinase C"
FT 658..660
FT /note= "O-phosphorylated by protein kinase C"
FT 723..725
FT /note= "O-phosphorylated by protein kinase C"
FT 767..769
FT /note= "O-phosphorylated by protein kinase C"
FT 817..819
FT /note= "O-phosphorylated by protein kinase C"
FT 866..868
FT /note= "O-phosphorylated by protein kinase C"
FT 1020..1022
FT /note= "O-phosphorylated by protein kinase C"
FT 1053..1055
FT /note= "O-phosphorylated by protein kinase C"
FT 1072..1074
FT /note= "O-phosphorylated by protein kinase C"
FT 1142..1144
FT /note= "O-phosphorylated by protein kinase C"
FT 1157..1159
FT /note= "O-phosphorylated by protein kinase C"
FT 1209..1211
FT /note= "O-phosphorylated by protein kinase C"
FT 1297..1299
FT /note= "O-phosphorylated by protein kinase C"
FT 1358..1360
FT Modified-site

FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT		37..40	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		44..47	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		110..113	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		121..124	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		256..259	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		444..447	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		640..643	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		693..696	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		739..742	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		756..759	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		817..820	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		824..827	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		887..890	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		1084..1087	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		1257..1260	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		1297..1300	
FT	Modified-site	/note= "O-phosphorylated by casein kinase III"	
FT		14..19	
FT	Modified-site	/note= "N-myristoylated"	
FT		20..25	
FT	Modified-site	/note= "N-myristoylated"	
FT		145..150	
FT	Modified-site	/note= "N-myristoylated"	
FT		170..175	
FT	Modified-site	/note= "N-myristoylated"	
FT		202..207	
FT	Modified-site	/note= "N-myristoylated"	
FT		260..265	
FT	Modified-site	/note= "N-myristoylated"	
FT		393..398	
FT	Modified-site	/note= "N-myristoylated"	
FT		418..423	
FT	Modified-site	/note= "N-myristoylated"	
FT		467..472	
FT	Modified-site	/note= "N-myristoylated"	
FT		515..520	
FT	Modified-site	/note= "N-myristoylated"	
FT		522..527	
Query Match 96.0%; Score 6791; DB 23; Length 1360;			
Best Local Similarity 95.8%; Pred. No. 0;			
Matches 1337; Conservative 3; Mismatches 7; Indels 48; Gaps 3;			
Qy	1	MTRKRTYWVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLDGPPWSQQERNPEAPGRAAVPP	60
Db	1	MTRKRTYWVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLDGPPWSQQERNPEAPGRAAVPP	60
	61	WGKYDAALRTMIPFRPKPRFPAPQPLDNAGLFSYLTVSWLTPLMIQLSRSLDENTIPPL	120
Db	61	WGKYDAALRTMIPFRPKPRFPAPQPLDNAGLFSYLTVSWLTPLMIQLSRSLDENTIPPL	120
Qy	121	SVHDASDKNVQRLHRLWEEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG	180
Db	121	SVHDASDKNVQRLHRLWEEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG	180
Qy	181	PILIIPKILEYSEEQLGNVVHVGGLCFALFLSECVKSLSFSSSWIINQRTAIRRAAVSS	240
Qy			

Db	181	PILIIPKILEYSEEQLGNVVHVGGLCFALFLSECVKSLSFSSSWIINQRTAIRRAAVSS	240
Qy	241	FAFEKLIQFKSVIHITSGE-----AISFFTGDVNVLFEGVCYGPLVLITCAS	287
Db	241	FAFEKLIQFKSVIHITSGE-----AISFFTGDVNVLFEGVCYGPLVLITCAS	300
	288	LVICSISYFIIGYTAFIAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQIRVTSEVLT	347
Db	301	LVICSISYFIIGYTAFIAILCYLLVFPPLAVFMTMAVKAQHHTSEVSDQIRVTSEVLT	360
Qy	348	CIKLIKMYTWEKPFAKIIEDLRRKERKLLKCKGLVQSLTSITLFIPTVATAVWVLIHTS	407
Db	361	CIKLIKMYTWEKPFAKIIEGM-----ESLTFCS-----	388
Qy	408	LKCLKTASMAFSMLASLNLRLSVFFVPPIAVKGLTNSKSAMRPFKKFFLQESPVFYVQTL	467
Db	389	---KPGDGMASFMLASLNLRLSVFFVPPIAVKGLTNSKSAMRPFKKFFLQESPVFYVQTL	445
Qy	468	QDPSKALVFEEATLSWQOTCPGIVNGALELERNGHASEGMTPRPRDALGPREEGNSLGP	527
Db	446	QDPSKALVFEEATLSWQOTCPGIVNGALELERNGHASEGMTPRPRDALGPREEGNSLGP	505
Qy	528	HKINLVVSKGMMLGVCNTGSGKSSLLSAILEEMHLLGSGVQGSLAYVPPQAWIVSGN	587
Db	506	HKINLVVSKGMMLGVCNTGSGKSSLLSAILEEMHLLGSGVQGSLAYVPPQAWIVSGN	565
Qy	588	IRENITLMGGAYDKARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRI	647
Db	566	IRENITLMGGAYDKARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRI	625
Qy	648	VYSRQIYLLDDPLSAVDAHVGKHIFEECIKKTLRGKTVVLVTHQLQYLEFCGQI	707
Db	626	VYSRQIYLLDDPLSAVDAHVGKHIFEECIKKTLRGKTVVLVTHQLQYLEFCGQI	685
Qy	708	GKICENGTHSELMQKKGKYAQLIQKMHKEATSDMLQDTAKIAEKPKVESQALATS	767
Db	686	GKICENGTHSELMQKKGKYAQLIQKMHKEATSDMLQDTAKIAEKPKVESQALATS	745
Qy	768	NGNAVPEHOLTQEEEMEESLSWRVYHHYIQAAGGYMVSCIIFFFVWLIVFLTIFS	827
Db	746	NGNAVPEHOLTQEEEMEESLSWRVYHHYIQAAGGYMVSCIIFFFVWLIVFLTIFS	805
Qy	828	SYWLEQSGSGTNSRESNGTMDLGNADNPQLSFYQLVYGLNALLICVGVCSG	887
Db	806	SYWLEQSGSGTNSRESNGTMDLGNADNPQLSFYQLVYGLNALLICVGVCSG	865
Qy	888	TRKASTALHNKLFNKVFERCPMSFFDTIPIGRLLNCFAGDLEQLDQLLP	947
Db	866	TRKASTALHNKLFNKVFERCPMSFFDTIPIGRLLNCFAGDLEQLDQLLP	925
Qy	948	VIAVLLIVSVLSPYILLMGAIMVICFIYYMMFKKAIGVFKRLENYSRSP	1007
Db	926	VIAVLLIVSVLSPYILLMGAIMVICFIYYMMFKKAIGVFKRLENYSRSP	985
Qy	1008	GLSSIHVYGKTEDFISQFKRLTDAQNYYLLFLSSSTRWMALRLEIMTNLVT	1067
Db	986	GLSSIHVYGKTEDFISQFKRLTDAQNYYLLFLSSSTRWMALRLEIMTNLVT	1045
Qy	1068	GISSTPYSFKVMVAVNIVQLASSFQATARIGLETEAQFAVERILQYMKMCV	1127
Db	1046	GISSTPYSFKVMVAVNIVQLASSFQATARIGLETEAQFAVERILQYMKMCV	1105
Qy	1128	GTSCPQGWPOHGEIIFQDYHMKYRDNTPTVLHGINLTIRGHEVVGVIRTS	1187
Db	1106	GTSCPQGWPOHGEIIFQDYHMKYRDNTPTVLHGINLTIRGHEVVGVIRTS	1165
Qy	1188	LFRLVEPMAGRILIDGVDICSIGLEDLRSKLSVIPQDPVLLSGTIRFNLD	1247
Db	1166	LFRLVEPMAGRILIDGVDICSIGLEDLRSKLSVIPQDPVLLSGTIRFNLD	1225
Qy	1248	WDALERTFLTKAISKFPKKLHTDVVENGNGNFSVGERQLLCIARAVLRNSKI	1307
Db	1226	WDALERTFLTKAISKFPKKLHTDVVENGNGNFSVGERQLLCIARAVLRNSKI	1285

QY 1308 IDMETDTLIQRTIREAFQGCTVLVIAHRVTTVLNCDHILVMGNGKVVEFDRPEVLRKKPG 1367
|||||
Db 1286 IDMETDTLIQRTIREAFQGCTVLVIAHRVTTVLNCDHILVMGNGKVVEFDRPEVLRKKPG 1345
|||||
QY 1368 SLFAALMATATSSLR 1382
|||||
Db 1346 SLFAALMATATSSLR 1360
|||||

RESULT 4
AAE22722
ID AAE22722 standard; Protein; 1331 AA.
XX
AC AAE22722;
XX
DT 09-AUG-2002 (first entry)
XX
DE Human adenosine triphosphate-binding cassette protein.
XX
KW Human; adenosine triphosphate-binding cassette; chemoprotectant; ATP;
KW ABC protein; gene therapy; breast proliferative fibrocystic disease;
KW breast adenocarcinoma; drug resistance; breast disorder; transgenic;
KW vaccine; ABCP.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 112..380
FT /note= "ABC transporter transmembrane domain"
FT Domain 114..133
FT /note= "Transmembrane domain"
FT Domain 247..269
FT /note= "Transmembrane domain"
FT Domain 337..353
FT /note= "Transmembrane domain"
FT Region 450..475
FT /note= "Antigenic epitope"
FT Region 486..657
FT /note= "ABCP biologically active portion; ABC
FT transporter domain"
FT Domain 493..500
FT /note= "ABC transporter walker site A motif; ATP-
FT binding site motif A (P-loop)"
FT Region 583..597
FT /note= "ABC transporter family signature"
FT Domain 603..607
FT /note= "ABC transporter walker site B motif"
FT Domain 750..774
FT /note= "Transmembrane domain"
FT Domain 755..1044
FT /note= "ABC transporter transmembrane domain"
FT Domain 807..832
FT /note= "Transmembrane domain"
FT Domain 891..912
FT /note= "Transmembrane domain"
FT Domain 909..928
FT /note= "Transmembrane domain"
FT Peptide 1002..1021
FT /label= Signal peptide
FT Region 1025..1040
FT /note= "Oligopeptide"
FT Domain 1117..1300
FT /note= "ABC transporter domain"
FT Domain 1124..1131
FT /note= "ABC transporter walker site A motif; ATP-
FT binding site motif A (P-loop)"
FT Region 1227..1241
FT /note= "ABC transporter family signature"
FT Domain 1247..1251
FT /note= "ABC transporter walker site B motif"
XX
PN W0200224742-A2.

XX 28-MAR-2002.
PD
XX
PF 20-SEP-2001; 2001WO-US29455.
XX
PR 22-SEP-2000; 2000US-0668628.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Krasnow RE, Baughn MR;
XX
DR WPI; 2002-394129/42.
DR N-PSDB; AAD36023.
XX
PT Mammalian adenosine triphosphate-binding cassette protein, and encoding
PT cDNA, useful for diagnosis and treatment of breast disorders, e.g.
PT breast proliferative fibrocystic disease and breast adenocarcinoma -
XX
PS Claim 15; Fig 1; 80pp; English.
XX
CC The invention relates to adenosine triphosphate (ATP)-binding cassette
CC (ABC) proteins (ABCP) and nucleic acid molecules encoding such proteins.
CC Sequences of the invention are useful for the diagnosis and treatment of
CC breast disorders, particularly breast proliferative fibrocystic disease,
CC breast adenocarcinoma and drug resistance. Polynucleotides of the
CC invention are used to produce transgenic cell lines or organisms which
CC are model systems for human breast disorders. They are also used in gene
CC therapy. ABCP sequences are useful in screening assays of phagemid or
CC B-lymphocyte immunoglobulin library to identify antibodies having desired
CC specificity. They are also useful as chemoprotectants and as vaccines.
CC The present sequence is human ABC protein.
XX
SQ Sequence 1331 AA;

Query Match 95.1%; Score 6723.5; DB 23; Length 1331;
Best Local Similarity 96.0%; Pred. NO. 0;
Matches 1327; Conservative 0; Mismatches 4; Indels 51; Gaps 1;

QY 1 MTRKRTYWPNSSGGLVNRGIDIGDDMVSGLIYKTYTLDQGPWSQQERNPEAPGRAAVPP 60
|||||
Db 1 MTRKRTYWPNSSGGLVNRGIDIGDDMVSGLIY----- 33

QY 61 WGKYDAALRTMIPFRPKRFPAPQPLDNAGLFSYLTVSWLTPLMIQSLRSRLDENTIPPL 120
|||||
Db 34 -----PLDNAGLFSYLTVSWLTPLMIQSLRSRLDENTIPPL 69

QY 121 SVHDASDKNVQRLHRLWEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG 180
|||||
Db 70 SVHDASDKNVQRLHRLWEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG 129

QY 181 PILIIPKILEYSEEQLGNVHVGVGLCFALFLSECVKSLSFSSSWIINQRTAIRFRAAVSS 240
|||||
Db 130 PILIIPKILEYSEEQLGNVHVGVGLCFALFLSECVKSLSFSSSWIINQRTAIRFRAAVSS 189

QY 241 FAFEKLIQFKSVIHITSGEAISFFTGDVNYLFEVCYGPLVLITCASLVICSISSYFIIG 300
|||||
Db 190 FAFEKLIQFKSVIHITSGEVISFFTGDVNYLFEVCYGPLVLITCASLVICSISSYFIIG 249

QY 301 YTAFAIILCYLLVPFLAVFMTRMVAKQHHTSEVSDQIRVTSEVLTICLIKMYTWEKP 360
|||||
Db 250 YTAFAIILCYLLVPFLAVFMTRMVAKQHHTSEVSDQIRVTSEVLTICLIKMYTWEKP 309

QY 361 FAKIIEDLRKERRKLLKCGLVQSLTSTLFIPTVATAVWVLIHTSLKLTASMAFSM 420
|||||
Db 310 FAKIIEDLRKERRKLLKCGLVQSLTSTLFIPTVATAVWVLIHTSLKLTASMAFSM 369

QY 421 LASLNLRLSVFFVPIAVKGLTNSKSAVMRFKKFFLOESPVFYVQTLQDPSKALVFEEAT 480
|||||
Db 370 LASLNLRLSVFFVPIAVKGPTNSKSAVMRFKKFFLOESPVFYVQTLQDPSKALVFEEAT 429

QY 481 LSWQOTCPGIVNGALELERNGHASEGMRPRDALGP EEGNSLGP ELHKINLVVSKGMML 540
|||||
Db 430 LSWQOTCPGIVNGALELERNGHASEGMRPRDALGP EEGNSLGP ELHKINLVVSKGMML 489

Qy	541	GVCNTGSGKSSLLSAILEEMHLLGSGVQGSLAYVPOQAWIVSGNIRENILMGCAYDK	600
Db	490	GVCNTGSGKSSLLSAILEEMHLLGSGVQGSLAYVPOQAWIVSGNIRENILMGCAYDK	549
Qy	601	ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKORISLARAVYSDRQIYLLDDP	660
Db	550	ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKORISLARAVYSDRQIYLLDDP	609
Qy	661	LSAVDAHVGKHIFEECIKKTLRGKTVVLVTHQLQYLEFCGQIILLENGKICENGTHSELM	720
Db	610	LSAVDAHVGKHIFEECIKKTLRGKTVVLVTHQLQYLEFCGQIILLENGKICENGTHSELM	669
Qy	721	QKKGKYAQLIQKMHKEATSDMLQDTAKIAEKPKVESQALATSLEESLNGNAVPEHQLTQE	780
Db	670	QKKGKYAQLIQKMHKEATSDMLQDTAKIAEKPKVESQALATSLEESLNGNAVPEHQLTQE	729
Qy	781	EEMEEGSLSWRVYHHYIOAAGGYMVSCIIFFFVVLIVFLTIFSFWMWSYWLEQGSGTNSS	840
Db	730	EEMEEGSLSWRVYHHYIOAAGGYMVSCIIFFFVVLIVFLTIFSFWMWSYWLEQGSGTNSS	789
Qy	841	RESNGTMADLGNADNPQLSPYQLVYGLNALLICVGCSSGIFTKVTRKASTALHNKLF	900
Db	790	RESNGTMADLGNADNPQLSPYQLVYGLNALLICVGCSSGIFTKVTRKASTALHNKLF	849
Qy	901	NKVFCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPIFSEQFLVLSLMVIAVLLIVSVLSP	960
Db	850	NKVFCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPIFSEQFLVLSLMVIAVLLIVSVLSP	909
Qy	961	YILLMGAIIMVICFIYYMFKKAIGVFKRLNENYSRSPFSHILNSLQGLSSIHVYGKTED	1020
Db	910	YILLMGAIIMVICFIYYMFKKAIGVFKRLNENYSRSPFSHILNSLQGLSSIHVYGKTED	969
Qy	1021	FISQFKRLTDAQNNYLLLFSSSTRWMLRLEIMTNLVTLAVALFVAFGISSTPYSPFKVMA	1080
Db	970	FISQFKRLTDAQNNYLLLFSSSTRWMLRLEIMTNLVTLAVALFVAFGISSTPYSPFKVMA	1029
Qy	1081	VNIVLQLASSFQATARIGLETEAQFTAVERILQYMKMCVSEAPLHMEGTSCPGWPQHGE	1140
Db	1030	VNIVLQLASSFQATARIGLETEAQFTAVERILQYMKMCVSEAPLHMEGTSCPGWPQHGE	1089
Qy	1141	IIFQDYHMKYRDNTPTVLHGINLTIRGHEVVGIVGRTGSGKSSLGMALFRLVEPMAGRIL	1200
Db	1090	IIFQDYHMKYRDNTPTVLHGINLTIRGHEVVGIVGRTGSGKSSLGMALFRLVEPMAGRIL	1149
Qy	1201	IDGVDICSIGLEDLRKLSVIPQDPVLLSGTIRFNLDPPDRHTDQQIWDALERTFLTAKI	1260
Db	1150	IDGVDICSIGLEDLRKLSVIPQDPVLLSGTIRFNLDPPDRHTDQQIWDALERTFLTAKI	1209
Qy	1261	SKFPKKLHTDVVENGGNFSVGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTI	1320
Db	1210	SKFPKKLHTDVVENGGNFSVGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTI	1269
Qy	1321	REAFQGCTVLVIAHRVTTVLNCDHILVMGNGKVVEFDRPEVLRKKPGSLFAALMATATSS	1380
Db	1270	REAFQGCTVLVIAHRVTTVLNCDRILVMGNGKVVEFDRPEVLRKKPGSLFAALMATATSS	1329
Qy	1381	LR 1382	
Db	1330	LR 1331	
RESULT 5			
AAB62558			
ID	AAB62558	standard; Protein; 1216 AA.	
XX			
AC	AAB62558;		
XX			
DT	23-JUL-2001	(first entry)	
XX			
DE	Novel human transporter protein (NHP).		
XX			
KW	Novel human protein; transporter protein; NHP; therapeutic; diagnostic;		

KW		gene therapy; screening.	
XX			
OS		Homo sapiens.	
XX			
PH		Key	Location/Qualifiers
FT		Misc-difference 86..87	/note= "the corresponding DNA encodes 3 amino acid
FT			residues in between the above positions which
FT			is not indicated in the present sequence"
XX		WO200132706-A2.	
PN		10-MAY-2001.	
PD			
XX		31-OCT-2000; 2000WO-US29852.	
PF			
XX		02-NOV-1999; 99US-0163018.	
PR		(LEXI-) LEXICON GENETICS INC.	
XX			
PA			
XX			
PI		Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;	
PI		Sands AT;	
XX			
DR		WPI; 2001-343477/36.	
DR		N-PSDB; AAF83646.	
XX		Novel isolated human polynucleotide sequences encoding polypeptides	
PT		that share sequence similarity with mammalian multidrug resistance	
PT		proteins and cellular transporter proteins, useful as probe or primer	
PT		-	
XX			
PS		Disclosure; Page 55-57; 59pp; English.	
XX			
CC		The invention relates to novel human transporter proteins (NHP) and	
CC		polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences	
CC		can be used in conjunction with PCR to screen libraries, isolate clones	
CC		and prepare cloning and sequencing templates. The NHP oligonucleotides	
CC		can also be used as hybridization probes for screening libraries, for	
CC		assessing gene patterns and for preparing antisense nucleic acid	
CC		molecules. The NHP nucleotide sequences are also useful in screening	
CC		techniques for drugs which treats symptomatic or phenotypic	
CC		manifestations of perturbing the normal function of NHP in the body.	
CC		Sequences AAB62547-558 represent NHP sequences.	
XX			
SQ		Sequence 1216 AA;	
		Query Match 86.2%; Score 6096; DB 22; Length 1216;	
		Best Local Similarity 87.8%; Pred. No. 0;	
		Matches 1214; Conservative 1; Mismatches 1; Indels 166; Gaps 2;	
Qy	1	MTRKRTYWVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQOERNPEAPGRAAVPP	60
Db	1	MTRKRTYWVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQOERNPEAPGRAAVPP	60
Qy	61	WGKYDAALRTMIPFRPKPRPPAPQPLDNAGLFSYLTVSWLTPLMIQSLRSRLDENTIPPL	120
Db	61	WGKYDAALRTMIPFRPKPRPPAPQPL--GLFSYLTVSWLTPLMIQSLRSRLDENTIPPL	117
Qy	121	SVHDASDKNVQRLHRLWEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG	180
Db	118	SVHDASDKNVQRLHRLWEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG	177
Qy	181	PILIIPKILEYSEEQNGNVVHGVGLCFALFLSECVKSLSFSSSWIINQRTAIRFRAAVSS	240
Db	178	PILIIPKILEYSEEQNGNVVHGVGLCFALFLSECVKSLSFSSSWIINQRTAIRFQAAVSS	237
Qy	241	FAFEKLIQFKSVIHITSGEATISFFFTGDVNYLFEGVCYGPLVLITCASLVICSISYFIIG	300
Db	238	FAFEKLIQFKSVIHITSGEATISFFFTGDVNYLFEGVCYGPLVLITCASLVICSISYFIIG	297
Qy	301	YTAFIAILCYLLVFPPLAVETMRMAVKAQHHTSEVSDQRIRVTSEVLTICIKLIKMYTWEKP	360
Db	298	YTAFIAILCYLLVFPPLVETMRMAVKAQHHTSEVSDQRIRVTSEVLTICIKLIKMYTWEKP	357

Db 241 EMHLEGSVGQSLAYVPQAWIVSGNIRENILMGAYDKARYLQVLHCCSLNRDLELL 300

QY 620 PFGDMTEIGERGLNLGGQKORISLARAVYSDRQIYLLDDPLSAVDHVGHKHFEECIKK 679

Db 301 PFGDMTEIGERGLNLGGQKORISLARAVYSDRQIYLLDDPLSAVDHVGHKHFEECIKK 360

QY 680 TLRGKTVVLVTHQLQYLEFCGQIILLNCKICENGTHSELMQKGYAQLIQKMHKEATS 739

Db 361 TLRGKTVVLVTHQLQYLEFCGQIILLNCKICENGTHSELMQKGYAQLIQKMHKEATS 420

QY 740 DMLQDTAKIAEKPKVESQALATSLEESLNGNAVPEHQLTQEEEMEEGSLSWRVYHHYIOA 799

Db 421 DMLQDTAKIAEKPKVESQALATSLEESLNGNAVPEHQLTQEEEMEEGSLSWRVYHHYIOA 480

QY 800 AGGYMVSCIIFFVVLIVFLTIFSFWMLSYWLEQSGGTNSSRESNGTMADLGNADNPQL 859

Db 481 AGGYMVSCIIFFVVLIVFLTIFSFWMLSYWLEQSGGTNSSRESNGTMADLGNADNPQL 540

QY 860 SFYQLVYGLNALLLICVGCSSGIFTKVTRKASTALHNKLFNKVPRCPMSFFDTIPIGRL 919

Db 541 SFYQLVYGLNALLLICVGCSSGIFTKVTRKASTALHNKLFNKVPRCPMSFFDTIPIGRL 600

QY 920 LNCFAGDLEQDLLPIFSEQFLVLSLMVIAVLIVSVLSPYILLMGAIIMVICFIYYMM 979

Db 601 LNCFAGDLEQDLLPIFSEQFLVLSLMVIAVLIVSVLSPYILLMGAIIMVICFIYYMM 660

QY 980 FKKAIGVFEXRLENYSRSLPFSHILNSLOGLSSIHVYKGTEDFISQFKRLTDAQNYYLLLF 1039

Db 661 FKKAIGVFEXRLENYSRSLPFSHILNSLOGLSSIHVYKGTEDFISQFKRLTDAQNYYLLLF 720

QY 1040 LSSTRWMALRLEIMTNLVTLAVALFVAFGISSTPYSEFKVMVAVNIVLQLASSFQATARIGL 1099

Db 721 LSSTRWMALRLEIMTNLVTLAVALFVAFGISSTPYSEFKVMVAVNIVLQLASSFQATARIGL 780

QY 1100 ETEAQFTAVERILOYMKMCVSEAPLHMEGTSCPQGWPHQGEIIFQDYHMKYRDNTPTVLH 1159

Db 781 ETEAQFTAVERILOYMKMCVSEAPLHMEGTSCPQGWPHQGEIIFQDYHMKYRDNTPTVLH 840

QY 1160 GINTLIRGHEVVGIVGRTGSGKSSLGMAFLRLEPVMAGRILIDGVDICSIGLEDLRSKLS 1219

Db 841 GINTLIRGHEVVGIVGRTGSGKSSLGMAFLRLEPVMAGRILIDGVDICSIGLEDLRSKLS 900

QY 1220 VIPQDPVLLSGTIRFNLDPFDRHTDQIWDALERTFLTKAISKFPKKLHTDVVENGNGFS 1279

Db 901 VIPQDPVLLSGTIRFNLDPFDRHTDQIWDALERTFLTKAISKFPKKLHTDVVENGNGFS 960

QY 1280 VGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGTCLVIAHRVTTV 1339

Db 961 VGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGTCLVIAHRVTTV 1020

QY 1340 LNCDHILVMNGKVVEFDRPEVLRKKPGSLFAALMATATSSLR 1382

Db 1021 LNCDHILVMNGKVVEFDRPEVLRKKPGSLFAALMATATSSLR 1063

RESULT 7

AAB62554

ID AAB62554 standard; Protein; 1021 AA.

XX

AC AAB62554;

XX

DT 23-JUL-2001 (first entry)

XX

DE Novel human transporter protein (NHP).

XX

KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;

KW gene therapy; screening.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Misc-difference 86...87

FT /note= "the corresponding DNA encodes 3 amino acid

FT residues in between the above positions which

FT is not indicated in the present sequence"

XX

PN WO200132706-A2.

XX

PD 10-MAY-2001.

XX

PF 31-OCT-2000; 2000WO-US29852.

XX

PR 02-NOV-1999; 99US-0163018.

XX

PA (LEXI-) LEXICON GENETICS INC.

XX

PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;

PI Sands AT;

XX

DR WPI; 2001-343477/36.

DR N-PSDB; AAF83642.

XX

PT Novel isolated human polynucleotide sequences encoding polypeptides

PT that share sequence similarity with mammalian multidrug resistance

PT proteins and cellular transporter proteins, useful as probe or primer

PT

PS Disclosure; Page 42-44; 59pp; English.

XX

CC The invention relates to novel human transporter proteins (NHP) and

CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences

CC can be used in conjunction with PCR to screen libraries, isolate clones

CC and prepare cloning and sequencing templates. The NHP oligonucleotides

CC can also be used as hybridization probes for screening libraries, for

CC assessing gene patterns and for preparing antisense nucleic acid

CC molecules. The NHP nucleotide sequences are also useful in screening

CC techniques for drugs which treats symptomatic or phenotypic

CC manifestations of perturbing the normal function of NHP in the body.

CC Sequences AAB62547-558 represent NHP sequences.

XX

SQ Sequence 1021 AA;

Query Match 73.8%; Score 5218.5; DB 22; Length 1021;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 1019; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 MTRKRTYVWPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQOERNPEAPGRAAVPP 60

Db 1 MTRKRTYVWPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQOERNPEAPGRAAVPP 60

QY 61 WGKYDAALRTMIPFRPKPRFPAQPLDNAGLFSYLTVSWLTPLMIQSLRSRLDENTIPPL 120

Db 61 WGKYDAALRTMIPFRPKPRFPAQPL--GLFSYLTVSWLTPLMIQSLRSRLDENTIPPL 117

QY 121 SVHDASDKNVQRLHRLWEEEVSRRGIEKASVLLVMLRFQRTLIPFDALLGICFCIASVLG 180

Db 118 SVHDASDKNVQRLHRLWEEEVSRRGIEKASVLLVMLRFQRTLIPFDALLGICFCIASVLG 177

QY 181 PILIIPKILEYSEEQLGNVHVHVGGLCFALFLSECVKSLSFSSSWIINQRTAIRFRAAVSS 240

Db 178 PILIIPKILEYSEEQLGNVHVHVGGLCFALFLSECVKSLSFSSSWIINQRTAIRFOAAVSS 237

QY 241 FAFEKLIQFKSVIHITSGEAISFFTGDNVYLFEGVCYGPLVLITCASLVICSISYFIIG 300

Db 238 FAFEKLIQFKSVIHITSGEAISFFTGDNVYLFEGVCYGPLVLITCASLVICSISYFIIG 297

QY 301 YTAFTAILCYLLVFPPLAVFMTRMAVKAQHHTSEVSDQRIRVTSEVLTICLIKMYTWEKP 360

Db 298 YTAFTAILCYLLVFPPLAVFMTRMAVKAQHHTSEVSDQRIRVTSEVLTICLIKMYTWEKP 357

QY 361 FAKIIEDLRRKERKLLKCKGLVQSLTSTITLFIPTVATAVWVLIHTSLKLTASMAFSM 420

Db 358 FAKIIEDLRRKERKLLKCKGLVQSLTSTITLFIPTVATAVWVLIHTSLKLTASMAFSM 417

QY 421 LASNLRLLSVFFVPIAVKGLTNSKSAVMRFKKFFLQESPVFYVQTLQDPSKALVFEEAT 480

Db 418 LASLNLRLSVFFVPIAVKGLTNSKSAVMRFKFFLOESPVYVQTLDPSKALVFEAT 477
QY 481 LSWQOTCPGIVNGALELERNGHASEGMTRPRDALGPPEEGNSLGPPELHKINLVVSKGMML 540
Db 478 LSWQOTCPGIVNGALELERNGHASEGMTRPRDALGPPEEGNSLGPPELHKINLVVSKGMML 537
QY 541 GVCGNTSGKSSLLSAILEEMHLLGSGVQGSLAYVPQQAIVSGNIRENILMGGAYDK 600
Db 538 GVCGNTSGKSSLLSAILEEMHLLGSGVQGSLAYVPQQAIVSGNIRENILMGGAYDK 597
QY 601 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRI SARAVYS DRQIYLLDDP 660
Db 598 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRI SARAVYS DRQIYLLDDP 657
QY 661 LSAVDHAVGKHIFEICKKTLRGKTVVLVTHQLQYLEFCGQIILLENGKICENGTHSELM 720
Db 658 LSAVDHAVGKHIFEICKKTLRGKTVVLVTHQLQYLEFCGQIILLENGKICENGTHSELM 717
QY 721 QKKGKYAQLIQMHKEATSDMLQDTAKIAEKP KVESQALATSLEESLNGNAVPEHQLTQE 780
Db 718 QKKGKYAQLIQMHKEATSDMLQDTAKIAEKP KVESQALATSLEESLNGNAVPEHQLTQE 777
QY 781 EEMEEGSLSWRVYHHYIQAAGGYMVSCIIFFFVVLIVFLTIFSFWWLSYWLEQSGGTNSS 840
Db 778 EEMEEGSLSWRVYHHYIQAAGGYMVSCIIFFFVVLIVFLTIFSFWWLSYWLEQSGGTNSS 837
QY 841 RESNGTMADLGNITADNPQLSFYQLVYGLNALLICVGCSSGIFTKVKTRKASTALHNKLF 900
Db 838 RESNGTMADLGNITADNPQLSFYQLVYGLNALLICVGCSSGIFTKVKTRKASTALHNKLF 897
QY 901 NKVFRCPMSFFDTPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLLIVSVLSP 960
Db 898 NKVFRCPMSFFDTPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLLIVSVLSP 957
QY 961 YILLGAIIMVICFIYMMFKKAIGVFKRLENYSRSPFLSHILNSLOGLSSIHVYGKTED 1020
Db 958 YILLGAIIMVICFIYMMFKKAIGVFKRLENYSRSPFLSHILNSLOGLSSIHVYGKTED 1017
QY 1021 FISQ 1024
Db 1018 FISQ 1021

RESULT 8
AAB62553
ID AAB62553 standard; Protein; 975 AA.
XX
AC AAB62553;
XX
DT 23-JUL-2001 (first entry)
XX
DE Novel human transporter protein (NHP).
XX
KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; screening.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 86..87 /note= "the corresponding DNA encodes 3 amino acid
FT residues in between the above positions which
FT is not indicated in the present sequence"
XX
PN WO200132706-A2.
XX
PD 10-MAY-2001.
XX
PF 31-OCT-2000; 2000WO-US29852.
XX
PR 02-NOV-1999; 99US-0163018.
XX

PA (LEXI-) LEXICON GENETICS INC.
XX
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
Sands AT;
XX
DR WPI; 2001-343477/36.
DR N-PSDB; AAF83641.
XX
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX
PS Disclosure; Page 39-41; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAB62547-558 represent NHP sequences.
XX
SQ Sequence 975 AA;

Query Match . 70.5%; Score 4985.5; DB 22; Length 975;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 973; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 MTRKRITYWVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQQERNPEAPGRAAVPP 60
Db 1 MTRKRITYWVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWSQQERNPEAPGRAAVPP 60
QY 61 WGKYDAALRTMIPFRPKPRFPAPQLDNAGLFSYLTVSWLTPLMIQSRSLDENTIPPL 120
Db 61 WGKYDAALRTMIPFRPKPRFPAPQL---GLFSYLTVSWLTPLMIQSRSLDENTIPPL 117
QY 121 SVHDASDKNVQRLHRLWEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG 180
Db 118 SVHDASDKNVQRLHRLWEEVSRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG 177
QY 181 PILIIPKILEYSEELGNVHVHVGCLFALFLSECVKSLSFSSSWIINORTAIRFRAAVSS 240
Db 178 PILIIPKILEYSEELGNVHVHVGCLFALFLSECVKSLSFSSSWIINORTAIRFQAAVSS 237
QY 241 FAFEKLIQFKSVIHITSGEAISFFTGDNVLFEGVCYGPLVLITCASLVICSISSYFIIG 300
Db 238 FAFEKLIQFKSVIHITSGEAISFFTGDNVLFEGVCYGPLVLITCASLVICSISSYFIIG 297
QY 301 YTAFAIILCYLLVFPFLAVFMTMAVKAQHHTSEVSDQRI RVTSEVLTICIKLIKMYTWEKP 360
Db 298 YTAFAIILCYLLVFPFLVFMTRMAVKAQHHTSEVSDQRI RVTSEVLTICIKLIKMYTWEKP 357
QY 361 FAKIIEDLRRKERKLLLEKCGLVQSLTSITLFIIPTVATAVWVLIHTSLKCLKLTASMAFSM 420
Db 358 FAKIIEDLRRKERKLLLEKCGLVQSLTSITLFIIPTVATAVWVLIHTSLKCLKLTASMAFSM 417
QY 421 LASLNLRLSVFFVPIAVKGLTNSKSAVMRFKFFLOESPVYVQTLDPSKALVFEAT 480
Db 418 LASLNLRLSVFFVPIAVKGLTNSKSAVMRFKFFLOESPVYVQTLDPSKALVFEAT 477
QY 481 LSWQOTCPGIVNGALELERNGHASEGMTRPRDALGPPEEGNSLGPPELHKINLVVSKGMML 540
Db 478 LSWQOTCPGIVNGALELERNGHASEGMTRPRDALGPPEEGNSLGPPELHKINLVVSKGMML 537
QY 541 GVCGNTSGKSSLLSAILEEMHLLGSGVQGSLAYVPQQAIVSGNIRENILMGGAYDK 600
Db 538 GVCGNTSGKSSLLSAILEEMHLLGSGVQGSLAYVPQQAIVSGNIRENILMGGAYDK 597
QY 601 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRI SARAVYS DRQIYLLDDP 660

Db 598 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNSGGQKQISLARAVYSDRQIYLLDDP 657
QY 661 LSAVDAHVGKHIFEICKKTLRGKTVVLVTHQLQYLEFCGQIILLENGKICENGTHSELM 720
Db 658 LSAVDAHVGKHIFEICKKTLRGKTVVLVTHQLQYLEFCGQIILLENGKICENGTHSELM 717
QY 721 QKKGKYAQLIOKMHEATSDMLQDTAKIAEKPKVESQALATSLEESLNGNAVPEHQLTQE 780
Db 718 QKKGKYAQLIOKMHEATSDMLQDTAKIAEKPKVESQALATSLEESLNGNAVPEHQLTQE 777
QY 781 EEMEEGSLSWRVYHHYIQAAGGYMVSCIIFFVFVLIVFLTIFSFWWLSYWLEQSGGTNSS 840
Db 778 EEMEEGSLSWRVYHHYIQAAGGYMVSCIIFFVFVLIVFLTIFSFWWLSYWLEQSGGTNSS 837
QY 841 RESNGTMADLGNADNPOLSFYQLVYGLNALLLICVGCSSGIFTKVTRKASTALHNKLF 900
Db 838 RESNGTMADLGNADNPOLSFYQLVYGLNALLLICVGCSSGIFTKVTRKASTALHNKLF 897
QY 901 NKVFRCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLLIVSVLSP 960
Db 898 NKVFRCPMSFFDTIPIGRLLNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLLIVSVLSP 957
QY 961 YILLMGAIIMVICFIYYM 978
Db 958 YILLMGAIIMVICFIYYM 975

RESULT 9
AAB62552
ID AAB62552 standard; Protein; 900 AA.
XX AAB62552;
AC AAB62552;
XX
DT 23-JUL-2001 (first entry)
XX
DE Novel human transporter protein (NHP).
XX
KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; screening.
XX
OS Homo sapiens.
XX
PN WO200132706-A2.
XX
PD 10-MAY-2001.
XX
PF 31-OCT-2000; 2000WO-US29852.
XX
PR 02-NOV-1999; 99US-0163018.
XX
PA (LEXI-) LEXICON GENETICS INC.
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI; 2001-343477/36.
DR N-PSDB; AAF83640.
XX
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT -
XX
PS Disclosure; Page 36-38; 59pp; English.

XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening

CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAB62547-558 represent NHP sequences.
XX
SQ Sequence 900 AA;
Query Match 63.3%; Score 4475.5; DB 22; Length 900;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 900; Conservative 0; Mismatches 0; Indels 163; Gaps 1;
QY 320 MTRMAVKAQHHTSEVSDQIRVTSEVLTCTIKLIKMYTWEKPFAKIIEDLRRKERKLEKC 379
Db 1 MTRMAVKAQHHTSEVSDQIRVTSEVLTCTIKLIKMYTWEKPFAKIIEDLRRKERKLEKC 60
QY 380 GLVQSLTSITLFIIPTVATAVWVLIHTSLKCLKLTASMAFSMLASNLRLSVFFVPPIAVK 439
Db 61 GLVQSLTSITLFIIPTVATAVWVLIHTSLKCLKLTASMAFSMLASNLRLSVFFVPPIAVK 120
QY 440 GLTNSKSAVMRFKKFFLOESPVFYVQTLQDPSKALVFEETLSWQOTCPGIVNGALELER 499
Db 121 GLTNSKSAVMRFKKFFLOESPVFYVQTLQDPSKALVFEETLSWQOTCPGIVNGALELER 180
QY 500 NGHASEGMRPRDALGP EEGNSLGP ELHKINLVYSGMMGLVCGNTSGKSSLSAILE 559
Db 181 NGHASEGMRPRDALGP EEGNSLGP ELHKINLVYSGMMGLVCGNTSGKSSLSAILE 240
QY 560 EMHLEGSGVGQSLAYVPQAWIVSGNIRENINILMGAYDKARYLQVLHCCSLNRDLELL 619
Db 241 EMHLEGSGVGQSLAYVPQAWIVSGNIRENINILMGAYDKARYLQVLHCCSLNRDLELL 300
QY 620 PFGDMTEIGERGLNLSGGQKQISLARAVYSDRQIYLLDDPLSAVDAAHVKGKHIFEICKK 679
Db 301 PFGDMTEIGERGLNLSGGQKQISLARAVYSDRQIYLLDDPLSAVDAAHVKGKHIFEICKK 360
QY 680 TLRGKTVVLVTHQLQYLEFCGQIILLENGKICENGTHSELMOCKGYAQLIOKMHEATS 739
Db 361 TLRGKTVVLVTHQLQYLEFCGQIILLENGKICENGTHSELMOCKGYAQLIOKMHEATS 420
QY 740 DMLQDTAKIAEKPKVESQALATSLEESLNGNAVPEHQLTQEEEMEESLSWRVYHHYIQ 799
Db 421 ----- 420
QY 800 AGGYMVSCIIFFFVVLIVFLTIFSFWWLSYWLEQSGTNSRESNGTMADLGNADNPQL 859
Db 421 ----- 420
QY 860 SFYQLVYGLNALLLICVGCSSGIFTKVTRKASTALHNKLFNKVFRCPMSFFDTIPIGRL 919
Db 421 -----VFCPMSFFDTIPIGRL 437
QY 920 LNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLLIVSVLSPYILLMGAIIMVICFIYYMM 979
Db 438 LNCFAGDLEQLDQLLPFSEQFLVLSLMVIAVLLIVSVLSPYILLMGAIIMVICFIYYMM 497
QY 980 FKKAIGVFKRLENYSRSPLEFSHILNSLQGLSSIHVYGKTEDFISQFKRLTDAQNNYLLLF 1039
Db 498 FKKAIGVFKRLENYSRSPLEFSHILNSLQGLSSIHVYGKTEDFISQFKRLTDAQNNYLLLF 557
QY 1040 LSSTRWMALRLEIMTNLVTLAVALFVAFGISSTPYSPFKVMVAVNIVLQASSFOATARIGL 1099
Db 558 LSSTRWMALRLEIMTNLVTLAVALFVAFGISSTPYSPFKVMVAVNIVLQASSFOATARIGL 617
QY 1100 ETEAQFTAVERILOYMKMCVSEAPLHMEGTSCPGWPQHGEIIFQDYHMKYRDNTPTVLH 1159
Db 618 ETEAQFTAVERILOYMKMCVSEAPLHMEGTSCPGWPQHGEIIFQDYHMKYRDNTPTVLH 677
QY 1160 GINLTIRGHEVVGIVGRTSGKSSLSGMALFRLVEPMAGRILIDGVDICSIGLEDLRSKLS 1219
Db 678 GINLTIRGHEVVGIVGRTSGKSSLSGMALFRLVEPMAGRILIDGVDICSIGLEDLRSKLS 737
QY 1220 VIPQDPVLLSGTIRENLDPPFDRHTDQIWDALERTFLTAKISKFPKKLHTDVVENGNGFS 1279
Db 738 VIPQDPVLLSGTIRENLDPPFDRHTDQIWDALERTFLTAKISKFPKKLHTDVVENGNGFS 797

QY 1280 VGERQLCIARAVLRNSKIILIDEATASIDMETDTLIORTIREAFQCTVLVIAHRVTTV 1339
|||||
Db 798 VGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIORTIREAFQCTVLVIAHRVTTV 857
QY 1340 LNCDHILVMGNKGVVEFDRPEVLRKKPGSLFAALMATATSSLR 1382
|||||
Db 858 LNCDHILVMGNKGVVEFDRPEVLRKKPGSLFAALMATATSSLR 900

RESULT 10
AAB62557
ID AAB62557 standard; Protein; 858 AA.
XX
AC AAB62557;
XX
DT 23-JUL-2001 (first entry)
XX
DE Novel human transporter protein (NHP).
XX
KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; screening.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 86..87 /note= "the corresponding DNA encodes 3 amino acid
FT residues in between the above positions which
FT is not indicated in the present sequence"
XX
PN WO200132706-A2.
XX
PD 10-MAY-2001.
XX
PF 31-OCT-2000; 2000WO-US29852.
XX
PR 02-NOV-1999; 99US-0163018.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
XX WPI; 2001-343477/36..
DR N-PSDB; AAF83645.
XX
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX
PS Disclosure; Page 52-54; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAB62547-558 represent NHP sequences.
XX
SQ Sequence 858 AA;

Query Match 60.5%; Score 4280; DB 22; Length 858;
Best Local Similarity 83.6%; Pred. No. 0;
Matches 856; Conservative 1; Mismatches 1; Indels 166; Gaps 2;

QY 1 MTRKRITYVVPNSSGGLVNRGIDIGDDMVSGLIYKTYTTLQDGPWSQQERNPEAPGRAAVPP 60
|||||
QY 61 WKYDAALRTMIPFRPKPRFPAPQPLDNAGLFSYLTVSWLTPLMIQSLRSRLDENTIPPL 120
|||||
Db 61 WKYDAALRTMIPFRPKPRFPAPQPL--GLFSYLTVSWLTPLMIQSLRSRLDENTIPPL 117
QY 121 SVHDASDKNVQRLHRLWEEEVRRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG 180
|||||
Db 118 SVHDASDKNVQRLHRLWEEEVRRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLG 177
QY 181 PILIIPKILEYSEEQLGNVVHVGVLCPALFSECVKLSFSSSWIINQRTAIRFRAAVSS 240
|||||
Db 178 PILIIPKILEYSEEQLGNVVHVGVLCPALFSECVKLSFSSSWIINQRTAIRFQAAVSS 237
QY 241 FAFEKLIQFKSVIHITSGEAISFFFTGDVNYLFEGVCYGPLVLTICASLVICSISYFIIG 300
|||||
Db 238 FAFEKLIQFKSVIHITSGEAISFFFTGDVNYLFEGVCYGPLVLTICASLVICSISYFIIG 297
QY 301 YTAFAIALCYLLVFPPLAVFMTRMAVKAQHHTSEVSDQIRVTSEVLTICKLIKMYTWEKP 360
|||||
Db 298 YTAFAIALCYLLVFPPLAVFMTRMAVKAQHHTSEVSDQIRVTSEVLTICKLIKMYTWEKP 357
QY 361 FAKIIEDLRKERRKLLLEKCGLVQSLTSITLFIPTVATAVWVLIHTSLKLTASMAFSM 420
|||||
Db 358 FAKIIEDLRKERRKLLLEKCGLVQSLTSITLFIPTVATAVWVLIHTSLKLTASMAFSM 417
QY 421 LASLNLRLSVFFVPIAVKGLTNSKSAMVRFKFFLQESPVFYVQTLQDPSKALVFEEAT 480
|||||
Db 418 LASLNLRLSVFFVPIAVKGLTNSKSAMVRFKFFLQESPVFYVQTLQDPSKALVFEEAT 477
QY 481 LSWQQTCPGIVNGALELERNGHASGEMTRPRDALGPPEEGNSLGPDELHKINLVVSKGMML 540
|||||
Db 478 LSWQQTCPGIVNGALELERNGHASGEMTRPRDALGPPEEGNSLGPDELHKINLVVSKGMML 537
QY 541 GVCGNTSGKSSLLSAILEEMHLLGSGVGQSLAYVPQQAIVSGNIRENILMGGAYDK 600
|||||
Db 538 GVCGNTSGKSSLLSAILEEMHLLGSGVGQSLAYVPQQAIVSGNIRENILMGGAYDK 597
QY 601 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRISLARAVYSDRQIYLLDDP 660
|||||
Db 598 ARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRISLARAVYSDRQIYLLDDP 657
QY 661 LSAYDAHVGKHIFEECIKKTLRGKTVLVTHQLQYLEFCGQIILLENGKICENGTHSELM 720
|||||
Db 658 LSAYDAHVGKHIFEECIKKTLRGKTVLVTHQLQYLEFCGQIILLENGKICENGTHSELM 717
QY 721 QKKGKYAQLIQKMHKEATSDMLQDTAKIAEKPKVESQALATSLSESLNGNAVPEHQLTQE 780
|||||
Db 718 QKKGKYAQLIQKMHKEATS----- 736
QY 781 EEMEEGSLSWRVYHHYIQAAGGYMVSCIIFFVVLIVFLTIFSFMWLSYWLEQGSGTNSS 840
Db 737 ----- 736
QY 841 RESNGTMADLGNADNPQLSFYQLVYGLNALLLICVGCSSGIFTKVKRKAATALHNKLF 900
Db 737 ----- 736
QY 901 NKVFRCPMSFFDTPIIGRLLNCFAGDLEQDLQLLPIFSEQFLVLSLMAVIAVLIVSVLSP 960
|||||
Db 737 --VFRCPMSFFDTPIIGRLLNCFAGDLEQDLQLLPIFSEQFLVLSLMAVIAVLIVSVLSP 794
QY 961 YILLMGAIMVICFIYMMFKKAIGVKRLENYSRSLFSLHLSLQGLSSIHVYKTED 1020
|||||
Db 795 YILLMGAIMVICFIYMMFKKAIGVKRLENYSRSLFSLHLSLQGLSSIHVYKTED 854
QY 1021 FISQ 1024
|||||
Db 855 FISQ 858
RESULT 11
AAB62556

QY 1 MTRKRITYVVPNSSGGLVNRGIDIGDDMVSGLIYKTYTTLQDGPWSQQERNPEAPGRAAVPP 60
|||||

XX 31-OCT-2000; 2000WO-US29852.
PF
XX
PR 02-NOV-1999; 99US-0163018.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI; 2001-343477/36.
DR N-PSDB; AAF83636.
XX
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX
PS Disclosure; Page 28-29; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAB62547-558 represent NHP sequences.
XX
SQ Sequence 705 AA;

Query Match 50.9%; Score 3598; DB 22; Length 705;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 MTRMAVKAQHHTSEVSDQIRIRVTSEVLTICIKLIKMYTWKPFKAKIIEDLRRKERKLLKEC 379
Db 1 MTRMAVKAQHHTSEVSDQIRIRVTSEVLTICIKLIKMYTWKPFKAKIIEDLRRKERKLLKEC 60

QY 380 GLVQSLTSITLFIPTVATAVVWLIHTSLKCLKLTASMAFSLASLNLLRSLSVFFVPIAVK 439
Db 61 GLVQSLTSITLFIPTVATAVVWLIHTSLKCLKLTASMAFSLASLNLLRSLSVFFVPIAVK 120

QY 440 GLTNSKSAVMRKFFLQESPVFYVQTLQDPSKALVFEEATLSWQQTCPGIVNGALELER 499
Db 121 GLTNSKSAVMRKFFLQESPVFYVQTLQDPSKALVFEEATLSWQQTCPGIVNGALELER 180

QY 500 NGHASEGWMTRPRDALGP EEGNSLGP ELHKINLVVSKGMMGLVCNGTSGKSSLLSAILE 559
Db 181 NGHASEGWMTRPRDALGP EEGNSLGP ELHKINLVVSKGMMGLVCNGTSGKSSLLSAILE 240

QY 560 EMHLLGSGVQGSLAYVPQQAIVSGNIRENILMGAYDKARYLQVLHCCSLNRDLELL 619
Db 241 EMHLLGSGVQGSLAYVPQQAIVSGNIRENILMGAYDKARYLQVLHCCSLNRDLELL 300

QY 620 PFGDMTEIGERGLNLSGGQKQRI SARAVYSDROIYLLDDPLSAVDAAHVKGKHIFEECIKK 679
Db 301 PFGDMTEIGERGLNLSGGQKQRI SARAVYSDROIYLLDDPLSAVDAAHVKGKHIFEECIKK 360

QY 680 TLRGKTVVLVTHQLQYLEFCGQI ILLNGKICENGTHSELMOQKGYAQLIQMHKEATS 739
Db 361 TLRGKTVVLVTHQLQYLEFCGQI ILLNGKICENGTHSELMOQKGYAQLIQMHKEATS 420

QY 740 DMLQDTAKIAEKP KVESQALATSLEESLNGNAVPEHQLTQEEEMEGSLSWRVYHHYIOA 799
Db 421 DMLQDTAKIAEKP KVESQALATSLEESLNGNAVPEHQLTQEEEMEGSLSWRVYHHYIOA 480

QY 800 AGGYMVSCIIFFFVWLIVFLTIFSFWWLSYWLEQSGSGTNSRESNGTMADLGNADNPQL 859
Db 481 AGGYMVSCIIFFFVWLIVFLTIFSFWWLSYWLEQSGSGTNSRESNGTMADLGNADNPQL 540

QY 860 SFYQLVYGLNALLICVGVCSGGIFTKVKTRKASTALHNKLFNKVFCPMSFFDTIPIGRL 919
Db
541 SFYQLVYGLNALLICVGVCSGGIFTKVKTRKASTALHNKLFNKVFCPMSFFDTIPIGRL 600

QY 920 LNCFAGDLEQLDQLLPFSEQFLVLSLMLVIAVLLIVSVLSPYILLMGALIMVICFIYYMM 979
Db
601 LNCFAGDLEQLDQLLPFSEQFLVLSLMLVIAVLLIVSVLSPYILLMGALIMVICFIYYMM 660

QY 980 FKKAIGVFKRLENYRSRPLFSHILNSLOGLSSIHVYGKTEDFISQ 1024
Db
661 FKKAIGVFKRLENYRSRPLFSHILNSLOGLSSIHVYGKTEDFISQ 705

RESULT 13
AAB62547
ID AAB62547 standard; Protein; 659 AA.
XX
AC AAB62547;
XX
DT 23-JUL-2001 (first entry)
XX
DE Novel human transporter protein (NHP).
XX
KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; screening.
XX
OS Homo sapiens.
XX
PN WO200132706-A2.
XX
PD 10-MAY-2001.
XX
PF 31-OCT-2000; 2000WO-US29852.
XX
PR 02-NOV-1999; 99US-0163018.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
WPI; 2001-343477/36.
DR N-PSDB; AAF83635.
XX
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX
PS Claim 3; Page 26-27; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAB62547-558 represent NHP sequences.
XX
SQ Sequence 659 AA;

Query Match 47.6%; Score 3365; DB 22; Length 659;
Best Local Similarity 100.0%; Pred. No. 3.8e-318;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 320 MTRMAVKAQHHTSEVSDQIRIRVTSEVLTICIKLIKMYTWKPFKAKIIEDLRRKERKLLKEC 379
Db 1 MTRMAVKAQHHTSEVSDQIRIRVTSEVLTICIKLIKMYTWKPFKAKIIEDLRRKERKLLKEC 60

QY 380 GLVQSLTSITLFIPTVATAVVWLIHTSLKCLKLTASMAFSLASLNLLRSLSVFFVPIAVK 439

Db 61 GLVQSLTSITLFIPTVATAVWVLHTSLKLTASMAFSLASNLRLSLSVFFVPIAVK 120
QY 440 GLTNSKSAVMRKKFFLQESPVYVQTLDPSKALVFEEATLSWQQTCPGIVNGALELER 499
Db 121 GLTNSKSAVMRKKFFLQESPVYVQTLDPSKALVFEEATLSWQQTCPGIVNGALELER 180
QY 500 NGHASEGMRPRDALGP EEGNSLGP ELHKINLVVSKGMLGVC GNTSGSKSSL SAILE 559
Db 181 NGHASEGMRPRDALGP EEGNSLGP ELHKINLVVSKGMLGVC GNTSGSKSSL SAILE 240
QY 560 EMHLLGSGVGQSLAYVPOQAWIVSGNIREN ILMGGAYDKARYLQVLHCCSLNRDLELL 619
Db 241 EMHLLGSGVGQSLAYVPOQAWIVSGNIREN ILMGGAYDKARYLQVLHCCSLNRDLELL 300
QY 620 PFGDMTEIGERGLNLSGGQKQRI SLARAVYSDRQIYLLDDPLSAVD AHVGKHIFEECIKK 679
Db 301 PFGDMTEIGERGLNLSGGQKQRI SLARAVYSDRQIYLLDDPLSAVD AHVGKHIFEECIKK 360
QY 680 TLRGKTVVLVTHQLQYLEFCGQIILL ENGKICENGTHSEL MQKKGKYAQLIQMHKEATS 739
Db 361 TLRGKTVVLVTHQLQYLEFCGQIILL ENGKICENGTHSEL MQKKGKYAQLIQMHKEATS 420
QY 740 DMLQDTAKIAEKP KVESQALATSL EESLNGNAVPEHQLTQEEEME EGSLSWRVYHHYIQA 799
Db 421 DMLQDTAKIAEKP KVESQALATSL EESLNGNAVPEHQLTQEEEME EGSLSWRVYHHYIQA 480
QY 800 AGGYMVSCIIFFFVVLIVELTFISFWWLSYWL EQSGTNSRESNGTMADLGN IADNPQL 859
Db 481 AGGYMVSCIIFFFVVLIVELTFISFWWLSYWL EQSGTNSRESNGTMADLGN IADNPQL 540
QY 860 SFYQLVYGLNALLLICVGVCSGIFTKVTRKASTALHNKLFNKVFCRCPMSFFDTIPIGR L 919
Db 541 SFYQLVYGLNALLLICVGVCSGIFTKVTRKASTALHNKLFNKVFCRCPMSFFDTIPIGR L 600
QY 920 LNCFAGDLEQLDQLLP IFSQFLVLSLMVIAVLLIVSVLSPYILLMGAIIMVICFIYYM 978
Db 601 LNCFAGDLEQLDQLLP IFSQFLVLSLMVIAVLLIVSVLSPYILLMGAIIMVICFIYYM 659

RESULT 14
AAE08075
ID AAE08075 standard; Protein; 1363 AA.
XX
AC AAE08075;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human transporter-related protein #22.
XX
KW Human; transporter-related protein; metabolite transporter;
KW organic cation transporter; multi-drug resistance; MDR; gene therapy;
KW cellular transporter; sodium-glucose cotransporter; diagnosis; screening;
KW symptomatic; phenotypic manifestation; biological condition.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1009..1010
FT /note= "Encoded by ACCTAGTTTAAAGACG"
FT Misc-difference 1121
FT /label= Unknown
FT /note= "Encoded by YGT"
XX
PN WO200157214-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US03646.
XX
PR 03-FEB-2000; 2000US-0179973.
PR 14-FEB-2000; 2000US-0182422.
XX

PA (LEXI-) LEXICON GENETICS INC.
XX
PI Turner CA, Mathur B, Wang X, Abuin A, Friedrich G, Zambrowicz B;
Sands AT;
XX
DR WPI; 2001-514599/56.
DR N-PSDB; AADI4909.
XX
PT Novel polynucleotides encoding novel human proteins with structural
PT similarity to cellular transporters for the diagnosis of disease and
PT use in gene therapy -
XX
PS Claim 4; Page 61-64; 91pp; English.
XX
CC The present sequence is a human transporter-related protein.
CC The human transporter-related protein share structural similarity with
CC mammalian metabolite or organic cation transporters, multi-drug
CC resistance (MDR) proteins, mammalian sodium-glucose cotransporters and
CC other cellular transporters. The transporter-related protein DNA may be
CC used for the detection of mutant sequences or inappropriately expressed
CC sequences for the diagnosis of disease. They may also be used to screen
CC for drugs effective in the treatment of the symptomatic or phenotypic
CC manifestations of perturbing the normal functions of the sequences of the
CC invention in the body. They may also be used in gene therapy for treating
CC biological conditions.
XX
SQ Sequence 1363 AA;
Query Match 47.3%; Score 3343; DB 22; Length 1363;
Best Local Similarity 48.7%; Pred. No. 2.4e-315;
Matches 653; Conservative 244; Mismatches 410; Indels 34; Gaps 6;
QY 63 KYDAALRTMIPERPKPRFPAPQPLDNAGLFSYLT VSWLTPLMTQSLRSRLDENTIPPLSV 122
Db 24 RYDPSLKTMI PVRPCARL-APNPVDDAGLLSFATFSWLT PVMVKG YRORLTVDTLPPLST 82
QY 123 HDASDKNVQRLHRLWEEVSRGIEKASVLLVMLRFQRTRLIPDALLGICFCIASVLGPI 182
Db 83 YDSSDTNAKRFRVLWDEEVARVCGPEKASLSHVVKFQRTRVLMDIVANILCIIMAAIGPT 142
QY 183 LIIPKILEYSEEQNGNVVHGVLGCFALFLSECVKSLSPSSSWIINQRTAIRFAAVSSFA 242
Db 143 VLIHOILQOQTER TSGKVWVGIGLCIALFATEFTKVFFWALAWAINYKTAIRLKVALSTLV 202
QY 243 FEKLIOFKSVIHITSGEAISFFTGDNVYLFE GVCYGPLVLITCASLVICSISSYFIIGYT 302
Db 203 FENLVSEKTLTHISVGEVLNLTSSDSYSLFEAALFCPLPATIPILMVFC AAYAFFILGPT 262
QY 303 AFIAILCYLLVFP LAVFMTRMAVKAQHHTSEVSDQRI RVTSVLTICIKLIKMYTWEKPFA 362
Db 263 ALIGISVYVIFIPVQMFMAKLSAFRRSAI LVTDKRVQTMNEFLTICIRLIKMYAWEKSFT 322
QY 363 KIIEDLRKKERK LLEKCGLVQSLTSITLFIPTVATAVWVLHTSLK LKLTASMAFSLA 422
Db 323 NTIQDIRRRERK LLEKAGFVQSGNSALAPIVSTIAIVLTLSCHILLRRKLTAPVAFSVIA 382
QY 423 SLNLLRLSVFFVPIAVKGLTNSKSAVMRKKFFLQESPVYVQTLDPSKALVFEEATLS 482
Db 383 MFNVMKFSIAILPFSIKAMAEANVSLRRMKKILIDKSPSPSYITQPEDPDTVLLLANATLT 442
QY 483 WQ-----QTCPGIVNGALEL---ERNGHASEGMRPRDALGP EEGNSLGP ELHKINLV 533
Db 443 WEHEASRKSTPKKLQNQKRHLCKKQ RSEAYSERSPPAKGGATGP EEGSDSLKSVLSISFV 502
QY 534 VSKGMMLGVC GNTSGSKSSL SAILEEMHLLGSGVGQSLAYVPOQAWIVSGNIREN IL 593
Db 503 VRKGKILGICGNVSGKSSL AALLGQMQLQKGVAVNGT LAYVYVSSQAWIFHGNVRENIL 562
QY 594 MGGAYDKARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRI SLARAVYSDRQ 653
Db 563 FGEKYDHQRYOHTVRVCGLOKDLNLPYGDLT EIGERGLNLSGGQKQRI SLARAVYSDRQ 622
QY 654 IYLLDDPLSAVD AHVGKHIFEECIKKTLRGKTVVLVTHQLQYLEFCGQIILL ENGKICEN 713

Db 623 LYLDDPLSAVDHVGKHFEECIKKTLRGKTVVLVTHQLQFLESCDEVILLEGEICEK 682
QY 714 GTHSELMQKKGKQAQLIQKMH-----KEATSDMLQDTAKIAEKPKV 754
Db 683 GTHKELMEERGRIAKLIHNLRLGLQFKDPEHLYNAAMVEAFKESPAEREEDAVLAPGNEKD 742
QY 755 ESQALATSLEESLNGNAVPEHQLTQEEEMEGLSWRVYHHYIOAAGYMVVSCIIFFEVV 814
Db 743 EGKSEETGSE--FVDTKVPEHQLIQTESPOEGTVTKTYHTYIKASGGYLLSLFTVFLFL 800
QY 815 LIVFLTIFSFWLSYWLQGGSGTNSRESNGTMADLGNADNPQLSFYQLVYGLNALLI 874
Db 801 LMIGSAAFSNNWLGWLWDKGRMTCGPQGNRTMCEVGAVLADIGQHVVQWVYTASVMFML 860
QY 875 CVGVCSSGIFTKVKTRKASTALHNKLFNKVFRCPMSFFDTPIGRLLNCFAGDLEQLDQLL 934
Db 861 VEGVTKGFEVTKTTLMASSSLHDTVFDKILKSPMSFFDTTPTGRLMNRFSDKMDEL DVRL 920
QY 935 PIFSEQFLVLSLMVIAVLIVSVLSPYILLMGAIIIMVICFIYYMMFKKAIGVFKRLENYS 994
Db 921 PFHAENFLQOFFMVVILVILAAVFPVAVLLVASLAVGFFILLRIFHRGVQELKKVENVS 980
QY 995 RSPLFSHILNSLOGLSSIHVYGKTEDFISQFKRLTDAONNYLLLLFLSSTRWMAALREIMT 1054
Db 981 RSPWFTHITSSMOGLGIIHAYGKKESCIT---TLNDENSSHLLYFNCALRWFAALRMDVLM 1037
QY 1055 NLVTLAVALFVAFGISSTPYSEFKVMVAVNIVLQLASSFQATARIGLETEAQFTAVERILQY 1114
Db 1038 NILFTVALLVTLFSFSSISTSSKGLSLSYIIQLSGLLQVCVRTGTETQAKFTSVELLREY 1097
QY 1115 MKMCVSEAPLHMEGTSCPPQGWPHGEIIFQDYHMKYRDNTPTVLHGINLTIIRGHEVVGIV 1174
Db 1098 ISTCVPECTHPLKVGTCPKDWPSEXGEITFRDYQMYRDNTPLVLDLSLNLNIQSGQTVGIV 1157
QY 1175 GRTSGKSSGLMALFRLVEPVMAGRILIDGVDICSIGLEDLRSKLSVIPODPVLLSGTIRF 1234
Db 1158 GRTSGKSSGLMALFRLVEPASGTIFIDEVDICLSLEDLRTKLTVIPODPVLFVGTVRY 1217
QY 1235 NLDPFDRHTDQIWDALERTFLTKAISKFPKKLHTDVVENGNGNFSVGERQLLCIARAVLR 1294
Db 1218 NLDPFESHTDEMLMQVLERTEFMRDITMKLPKLOAEVTENGENFSVGERQLLCVARALLR 1277
QY 1295 NSKIILIDEATASIDMETDTLIQRTIREAFQCGCTVLVIAHRVTTVLNCDHILVMNGKVV 1354
Db 1278 NSKIILLDEATASMSKDTLTVQNTIKDAFKGCTVLTIAHRLNTVLNCDHVLVMENGKVI 1337
QY 1355 EFDRPEVLRKKPGSLFAALMA 1375
Db 1338 EFDKPEVLAEKPDSAFAMLLA 1358

RESULT 15
AAE08078
ID AAE08078 standard; Protein; 1332 AA.
XX
AC AAE08078;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human transporter-related protein #25.
XX
KW Human; transporter-related protein; metabolite transporter;
KW organic cation transporter; multi-drug resistance; MDR; gene therapy;
KW cellular transporter; sodium-glucose cotransporter; diagnosis; screening;
KW symptomatic; phenotypic manifestation; biological condition.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1009..1010
FT /note= "Encoded by ACCTAGTTTAAGACG"
FT Misc-difference 1121

FT /label= Unknown
FT /note= "Encoded by YGT"
XX
PN WO200157214-A2.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-US03646.
PR 03-FEB-2000; 2000US-0179973.
PR 14-FEB-2000; 2000US-0182422.
XX (LEXI-) LEXICON GENETICS INC.
PA
XX Turner CA, Mathur B, Wang X, Abuin A, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI; 2001-514599/56.
DR N-PSDB; AAD14912.
XX
PT Novel polynucleotides encoding novel human proteins with structural
PT similarity to cellular transporters for the diagnosis of disease and
PT use in gene therapy -
XX
PS Disclosure; Page 74-77; 9lpp; English.
XX
CC The present sequence is a human transporter-related protein.
CC The human transporter-related protein share structural similarity with
CC mammalian metabolite or organic cation transporters, multi-drug
CC resistance (MDR) proteins, mammalian sodium-glucose cotransporters and
CC other cellular transporters. The transporter-related protein DNA may be
CC used for the detection of mutant sequences or inappropriately expressed
CC sequences for the diagnosis of disease. They may also be used to screen
CC for drugs effective in the treatment of the symptomatic or phenotypic
CC manifestations of perturbing the normal functions of the sequences of the
CC invention in the body. They may also be used in gene therapy for treating
CC biological conditions.
XX
SQ Sequence 1332 AA;

Query Match 45.4%; Score 3208.5; DB 22; Length 1332;
Best Local Similarity 47.4%; Pred. No. 2.9e-302;
Matches 636; Conservative 238; Mismatches 402; Indels 65; Gaps 7;
QY 63 KYDAALRTMIPFRPKPRFPAPQPLDNAGLFSYLTVSWLTPLMIQLSRSLDENTIPPLSV 122
Db 24 RYDPSLKTMIPIVRPCARL-APNPVDDAGLLSFATFSWLTPVMVKGYRQLTVDTLPLST 82
QY 123 HDASDKNVQRLHRLWEEEVRRGIEKASVLLVMLRFQRTLIFDALLGICFCIASVLGPI 182
Db 83 YDSSDTNAKRFVLDVEVARVGPPEKASLSHVVKFQRTVRLMDIVANILCIIMAAIGPT 142
QY 183 LIIPKILEYSEEQLGNVVHVGVLGCLFALFELSECVKLSFSSSWIINQRTAIRFRAAVSSFA 242
Db 143 VLIHQILQQTERTSGKVWVGIGLCIALFATEFTKVFFWALAWAINRYTAIRLKVALSTLV 202
QY 243 FEKLIQKSVIHITSGEAISFFTGDVNYLFEQVCYGPLVLITCASLVICSISSYFIIGYT 302
Db 203 FENLVSFKTLTHISVGEVLNILSSDSYSLFEAALFCPLPATIPILMVFCAAYAFFILGPT 262
QY 303 AFIAILCYLLVFPFLAVFMTRMVAKQAHHITSEVSDQIRVTVSEVLTICIKLIKMYTWKPFA 362
Db 263 ALIGISVYVIFIPVQMFMAKLNSAFRRSAILVTDKRVQTMNEFLTCLIRLIKMYAWEKST 322
QY 363 KIIEDLRRKERKLLKCGLVQSLTSITLFIPTVATAVWVLIHTSLKILKTASMAFSMLA 422
Db 323 NTIQDIRRRERKLLKAGFVQSGNSALAPIVSTIAIVLTLSCHILLRRLKLTAPVAFSVIA 382
QY 423 SLNLLRLSVFVPIAVKGLTNSKSAVMRPFKFFLQESPVFYVQTLQDPSKALVFEETLS 482
Db 383 MFNVMKFSIAILPFSSIKAMAEANVSLRRMKKILIDKSPSYITQPEDPDTVLLLANATLT 442
QY 483 WQ-----QTCPGIVNGALEL---ERNCHASEGMRPRDALGPPEEGNSLGPELHKINLV 533

Db 443 WEHEASRKSTPKKLQNKRRHLCKKQKRSAYSEKSPPAKGATGPEEQSDSLKSVLHSISFV 502
QY 534 VSKGMLGVCNTGSGKSSLLSAILEEMHLLGSGVGQSLAYVPQQAIVSGNIRENIL 593
Db 503 VRKGKILGICGNVSGKSSLLAALLQGMQKQGVAVNGTLAYVSQQAIFHGNVRENIL 562
QY 594 MGGAYDKARYLQVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQKORISLARAVYSDRQ 653
Db 563 FGEKYDHORYQHTVRVCGLQKDLNLPYGDLTEIGERGLNLSGGQKQKORISLARAVYSDRQ 622
QY 654 IYLLDDPLSAVDHAVGKHIFECEIKKTLRGKTVVLVTHQLYLEFCGQIILLENGKICEN 713
Db 623 LYLLDDPLSAVDHAVGKHVFEECIKKTLRGKTVVLVTHQLQFLESCDEVILLEDEICEK 682
QY 714 GTHSELMQKKGKYAQLIQKMH-----KEATSDMLQDHTAKIAEKPKV 754
Db 683 GTHKELMEERGRYAKLIHNLRLGLQKDPHEHLYNAAMVEAFKESPAEREEDAVLAPGNEKD 742
QY 755 ESQALATSLFEESLNGNAVPEHQLTQEEEMEESLSWRVYHHYIOAAGGYMVSCIIFFFV 814
Db 743 EGKESGTGSE--FVDTKVPEHQLIQTESPQEGTVTKTYHTYIKASGGYLLSLFTVFLFL 800
QY 815 LIVFLTIFSPFWLSYWLQSGSGTSSRESNGTMADLGNADNPQLSFYQLVYGLNALLLI 874
Db 801 LMIGSAAFSNWMLGLWLDKGSRTCCGPQGNRTMCEVGAVLADICQHVVYQWVYTASMFML 860
QY 875 CVGVCSSGIPTKVTRKASTALHNKLFNKVRCPMSEFDTPIGRLLNCFAGDLEQLDQLL 934
Db 861 VFGVTKGFEVTKTTLMASSSLHDTVEDKILKSPMSFDDTPTGRLMNRFSDKMDELVDRL 920
QY 935 PIFSEQFLVLSLVMIAVLLIVSVLSPYILLMGAIIMVICFIYMMFKKAIGVKRLENYS 994
Db 921 PFHAENLQOFFMVFILVILAAVFPVAVLLVVASLAVGFFILLRIFHRGVQELKKVENVS 980
QY 995 RSPLFSHILNSLQGLSSIHVYKTEDFISQFKRLTDAQNNYLLFLLSSTRWMALRLEIMT 1054
Db 981 RSPWFTHITSSMQGLGIIHAYGKKESCIT---TLNDENSSHLLYFNCALRWFALRMDVLM 1037
QY 1055 NLVTLAVALFVAFGISSTPYSEFKVMVNIIVLQASSFQATARIGLETEAQFTAVERILQY 1114
Db 1038 NILTFTVALLVTLSESSISTSSKGLSLSYIIQLSGLLQVCVRTCTETQAKFTSVELLREY 1097
QY 1115 MKMCVSEAPLHMEGTSCPOGWPOHGELIFQDYHMKYRDNTPTVLHGINTIRGHEVVGIV 1174
Db 1098 ISTCVPECTHPLKVGTCPKDWPSXGELTFRDYQMRYRDNTPLVLDSLNLNIQSGQTVGIV 1157
QY 1175 GRTGSGKSSLGMALFRLVEPMAGRILIDGVDICSIGLEDLRSKLSVIPQDPVLLSGTTRF 1234
Db 1158 GRTGSGKSSLGMALFRLVEPASGTIFIDEVDICILSLEDLRTKLTVIPQDPVLF----- 1211
QY 1235 NLDPFDRHTDQOIWDALERTFLTKAISKFPKKLHTDVVENGGNFSVGERQLLICIARAVLR 1294
Db 1212 -----IMKLPEKLQAEVTENGENFSVGERQLLCVARALLR 1246
QY 1295 NSKIIILIDEATASIDMETDTLIORTIREAFQGCTVLVIAHRVTTVLNCDHILVMGNKV 1354
Db 1247 NSKIIILDEATASMDSKTDTLVQNTIKDAFKGCTVLTIAHRLNTVLNCDHVLVMENGV 1306
QY 1355 EFDREPEVLRKKPGSLFAALMA 1375
Db 1307 EFDKPEVLAEKPDSAFAMLLA 1327

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2003, 05:07:50 ; Search time 946 Seconds
(without alignments)
11574.224 Million cell updates/sec

Title: US-10-087-782A-1
Perfect score: 4862
Sequence: 1 actgggataaagcaagaaga.....tataaaactaaggaaaactc 4862

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :					N_Geneseq_101002:*				
1:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1980.DAT:*						
2:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1981.DAT:*						
3:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1982.DAT:*						
4:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1983.DAT:*						
5:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1984.DAT:*						
6:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1985.DAT:*						
7:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1986.DAT:*						
8:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1987.DAT:*						
9:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1988.DAT:*						
10:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1989.DAT:*						
11:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1990.DAT:*						
12:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1991.DAT:*						
13:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1992.DAT:*						
14:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1993.DAT:*						
15:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1994.DAT:*						
16:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1995.DAT:*						
17:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1996.DAT:*						
18:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1997.DAT:*						
19:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1998.DAT:*						
20:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA1999.DAT:*						
21:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA2000.DAT:*						
22:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA2001A.DAT:*						
23:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA2001B.DAT:*						
24:	/SIDS2/gcgdata/	geneseq/	geneseqn-emb1/NA2002.DAT:*						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4391.4	90.3	4638	ABA92270	Human ATP-binding
2	4249.6	87.4	4427	ABK92211	Prostate cancer-as
3	4142.6	85.2	4149	AAF83643	Novel human transp
4	4133.6	85.0	5020	AAF83647	Novel human transp
5	3964.2	81.5	4074	AAD36023	Human adenosine tr
6	3185.8	65.5	3189	AAF83637	Novel human transp
7	3156.2	64.9	3660	AAF83646	Novel human transp
8	3066.6	63.1	3075	AAF83642	Novel human transp
9	2967.2	61.0	3055	ABV21036	Human prostate exp

10	2967.2	61.0	3055	23	ABV26879	Human prostate exp
11	2927.6	60.2	2937	22	AAF83641	Novel human transp
12	2213.2	45.5	2448	22	AAF83644	Novel human transp
13	2213.2	45.5	2586	22	AAF83645	Novel human transp
14	2199.4	45.2	2700	22	AAF83640	Novel human transp
15	2111.8	43.4	2115	22	AAF83636	Novel human transp
16	2102.2	43.2	2699	24	AAD33647	Human TRICH-2 cDNA
17	2010	41.3	2010	24	AAD36038	Human Adenosine tr
18	1973.8	40.6	1977	22	AAF83635	Novel human transp
19	1279.2	26.3	4101	22	AAD14909	Human transporter-
20	1259.4	25.9	1488	22	AAF83638	Novel human transp
21	1259.4	25.9	1626	22	AAF83639	Novel human transp
22	1132.8	23.3	4008	22	AAD14912	Human transporter-
23	1063	21.9	3660	22	AAD14910	Human transporter-
24	1057.4	21.7	3753	22	AAD14911	Human transporter-
25	768	15.8	4781	22	AAC85287	Multidrug-resistan
26	768	15.8	4847	19	AAV65682	Human multidrug re
27	768	15.8	4847	21	AAA40481	Human MRP-beta cDN
28	767.4	15.8	3549	23	AA590084	DNA encoding novel
29	764.8	15.7	5838	20	AAZ30079	cDNA encoding a hu
30	764.8	15.7	5838	21	AAZ94745	Human App binding
C	31	641.2	732	24	AAD36033	Human adenosine tr
	32	629.8	1698	22	AAD14913	Human transporter-
C	33	624	673	24	AAD36029	Human adenosine tr
	34	595.2	609	24	AAD36034	Human adenosine tr
C	35	591	591	24	AAD36024	Human adenosine tr
	36	566	578	24	AAD36025	Human adenosine tr
C	37	555	575	24	AAD36032	Human adenosine tr
	38	524.2	588	24	AAD36026	Human adenosine tr
C	39	506.8	521	24	AAD36028	Human adenosine tr
	40	506.4	571	24	AAD36027	Human adenosine tr
C	41	497.2	589	24	AAD36031	Human adenosine tr
	42	483.4	1605	22	AAD14916	Human transporter-
C	43	443.2	4864	17	AAT14911	cDNA encoding mult
	44	443.2	4885	17	AAT14910	cDNA encoding mult
45	443.2	9.1	5011	15	AAQ65377	Multidrug resistan

ALIGNMENTS

RESULT 1	
ABA92270	
ID	ABA92270 standard; cDNA; 4638 BP.
XX	
AC	ABA92270;
XX	
DT	10-JUN-2002 (first entry)
DE	Human ATP-binding cassette transporter 44589 cDNA.
XX	
KW	ATP-binding cassette transporter; ABC transporter 44589; human;
KW	cancer; osteopathic; immunomodulatory; cardiovascular; analgesic;
KW	antidiabetic; cytostatic; antileukaemic; hepatic; diagnosis;
KW	gene therapy; vaccine; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	201..4283
FT	/*tag= a
FT	/note= "the coding region is also specifically
FT	claimed in Claim 1"
XX	
PN	WO200216589-A2.
XX	
PD	28-FEB-2002.
XX	
PF	21-AUG-2001; 2001WO-US26096.
XX	
PR	21-AUG-2000; 2000US-226770P.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.

XX Curtis RAJ;
PI
XX
DR WPI; 2002-257915/30.
DR P-PSDB; AAM51155.
XX
PT Novel human ATP-binding cassette transporter family member (designated
PT 44589) proteins and nucleic acids, useful for preventing, diagnosing
PT and treating e.g. cancer, hepatic disorders and bone diseases -
XX
XX
PS Claim 1; Page 115-117; 133pp; English.
XX
CC The present sequence, the coding region of which is also claimed,
CC encodes a novel 1360-amino acid human ATP-binding cassette (ABC)
CC transporter family member, designated 44589 (see AAM51155). The
CC invention provides isolated 44589 nucleic acids (including
CC antisense nucleic acids), recombinant expression vectors containing
CC the nucleic acids, host cells into which the expression vectors
CC have been introduced, and nonhuman transgenic animals in which
CC the 44589 gene has been introduced or disrupted. It also provides
CC 44589 proteins, fusion proteins, antigenic peptides and anti-44589
CC antibodies. The nucleic acids and polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate 44589 expression, e.g. by rectifying mutations or
CC deletions in a patient's genome that affect the activity of 44589
CC by expressing inactive proteins or to supplement the patient's own
CC production of 44589. The nucleic acids may also be used as probes
CC in diagnostic assays. A method for identifying a compound that
CC binds 44589 polypeptide is claimed, and also a method of inhibiting
CC aberrant activity of a 44589-expressing cell, especially a cancer
CC cell, by contacting with a compound that reduces or inhibits 44589
CC activity of expression e.g. a peptide, phosphopeptide, small
CC organic molecule or antibody. Diseases that may be prevented,
CC diagnosed and treated include e.g. bone disorders, immune disorders,
CC cardiovascular disorders, pain, lipid metabolism, diabetes,
CC cancers and leukaemias, and liver disorders.
XX
SQ Sequence 4638 BP; 1106 A; 1172 C; 1223 G; 1137 T; 0 other;

Query Match 90.3%; Score 4391.4; DB 24; Length 4638;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 4578; Conservative 0; Mismatches 21; Indels 144; Gaps 3;

QY 151 GATGTTTAAAGAGGGATCAAGCACAGGCTAAGGAGAGGAAGACAGGACCCCAAACC 210
Db 1 GATGTTTAAAGAGGGATCAAGCACAGGCTAAGGAGAGGAAGACAGGACCCCAAACC 60
QY 211 TCTGCATGGCCCCAATATGCTCCCTGCGAGGGTAGTGCCCCCTCTCTGGCTGCTCAAGGC 270
Db 61 TCTGCATGGCCCCAATATGCTCCCTGCGAGGGTAGTGCCCCCTCTCTGGCTGCTCAAGGC 120
QY 271 GAGATCTAAGCTTCTTCTAAGCTTCTGCTGCTTTTCTCATATTTCTGATTTCTGGAAACGA 330
Db 121 GAGATCTAAGCTTCTTCTAAGCTTCTGCTGCTTTTCTCATATTTCTGATTTCTGGAAACGA 180
QY 331 AGAATTGGCAGGAACCTGAAAATGACTAGGAAGAGGACATACCTGGTGCCCAACTCTTCTG 390
Db 181 AGAATTGGCAGGAACCTGAAAATGACTAGGAAGAGGACATACCTGGTGCCCAACTCTTCTG 240
QY 391 GTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATA 450
Db 241 GTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATA 300
QY 451 AAACCTATACTCTCCAAGATGGCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAG 510
Db 301 AAACCTATACTCTCCAAGATGGCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAG 360
QY 511 GGAGGGCAGCTGTCCCACCGTGGGGAAGTATGATGCTGCCCTTGAGAACCATGATTCCT 570
Db 361 GGAGGGCAGCTGTCCCACCGTGGGGAAGTATGATGCTGCCCTTGAGAACCATGATTCCT 420
QY 571 TCCGTCCCAAGCCGAGGTTTCTTGCCCCCAGCCCCCTGGACAATGCTGGCCTGTCTCCT 630
Db 1 TCCGTCCCAAGCCGAGGTTTCTTGCCCCCAGCCCCCTGGACAATGCTGGCCTGTCTCCT 1455

Db 421 TCCGTCCCAAGCCGAGGTTTCTTGCCCCCAGCCCCCTGGACAATGCTGGCCTGTCTCCT 480
QY 631 ACCTCACCGTGTATGGCTCACCCCGCTCATGATCCAAAGCTTACGAGTTCGCTTAGATG 690
Db 481 ACCTCACCGTGTATGGCTCACCCCGCTCATGATCCAAAGCTTACGAGTTCGCTTAGATG 540
QY 691 AGAACACCATCCCTCCACTGTGTCAGTCCATGATGCCTCAGACAAAATGTCCAAAGGCTTC 750
Db 541 AGAACACCATCCCTCCACTGTGTCAGTCCATGATGCCTCAGACAAAATGTCCAAAGGCTTC 600
QY 751 ACCGCCCTTTGGGAAGAAGTCTCAAGCGGAGGGATTGAAAAGCTTCAGTGTCTCTGG 810
Db 601 ACCGCCCTTTGGGAAGAAGTCTCAAGCGGAGGGATTGAAAAGCTTCAGTGTCTCTGG 660
QY 811 TGATGTCAGGTTCCAGAGAACAAAGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTCT 870
Db 661 TGATGTCAGGTTCCAGAGAACAAAGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTCT 720
QY 871 GCATTGCCAGTGTACTCGGGCCAAATATTGATTATATACCAAAGATCCTGGAATATTCAGAAG 930
Db 721 GCATTGCCAGTGTACTCGGGCCAAATATTGATTATATACCAAAGATCCTGGAATATTCAGAAG 780
QY 931 AGCAGTTGGGGAATGTTGTCCATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAAT 990
Db 781 AGCAGTTGGGGAATGTTGTCCATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAAT 840
QY 991 GTGTGAAGTCTCTGAGTTTCTCCTCCAGTTGGATCATCAACCAACGACACAGCCATCAGGT 1050
Db 841 GTGTGAAGTCTCTGAGTTTCTCCTCCAGTTGGATCATCAACCAACGACACAGCCATCAGGT 900
QY 1051 TCCGAGCAGCTGTTTCTCCTTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATAC 1110
Db 901 TCCGAGCAGCTGTTTCTCCTTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATAC 960
QY 1111 ACATCACCTCAGGAG-----AGGCCA 1131
Db 961 ACATCACCTCAGGAGAGGGAGGTGACATCTGTGCCCATCAACTTGCTGTCTTGCAGGCCA 1020
QY 1132 TCAGCTTCTTCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGCTATGGACCCCTAG 1191
Db 1021 TCAGCTTCTTCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGCTATGGACCCCTAG 1080
QY 1192 TACTGATCACCTGCGCATCGTGGTCACTGCAGCATTTCTTCCACTTTCATTTGGAT 1251
Db 1081 TACTGATCACCTGCGCATCGTGGTCACTGCAGCATTTCTTCCACTTTCATTTGGAT 1140
QY 1252 ACATGCAATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCCTGCGGTATTCATGA 1311
Db 1141 ACATGCAATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCCTGCGGTATTCATGA 1200
QY 1312 CAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGAGGACCGCATCCGTG 1371
Db 1201 CAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGAGGACCGCATCCGTG 1260
QY 1372 TGACCAGTGAAGTTCTCACTTGCATTAAAGCTGATTAAAATGTACATGGGAGAAACCAT 1431
Db 1261 TGACCAGTGAAGTTCTCACTTGCATTAAAGCTGATTAAAATGTACATGGGAGAAACCAT 1320
QY 1432 TTGCAAAAATCATTGAAGACCTTAAGAAGGAAGAAAGAACTATTGGAGAAGTGCGGGC 1491
Db 1321 TTGCAAAAATCATTGAAG-----1338
QY 1492 TTGTCCAGAGCCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGGTCT 1551
Db 1339 -----GTATGGAAGTCTGACTTCTGCTCCAAACCTGGTGATGGC-----1379
QY 1552 GGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATGC 1611
Db 1380 -----ATGGCCTTCAGCATGC 1395
QY 1612 TGGCCTCCTTGAATCTCCTTCGGCTGTCAGTGTCTTTGTGCCCTATTGCAGTCAAAGGTC 1671
Db 1396 TGGCCTCCTTGAATCTCCTTCGGCTGTCAGTGTCTTTGTGCCCTATTGCAGTCAAAGGTC 1455

Db	3616	TC AACCTGACCATCCGCGCCACGAAGTGCTGGGCATCGTGGGAAGGACGGGCTCTGGGA	3675
QY	3892	AGTCCCTCCTTGGGCATGGCTCTCTTCCGCGCTGGTGGAGCCCATGGCAGCCGGATTCTCA	3951
Db	3676	AGTCCCTCCTTGGGCATGGCTCTCTTCCGCGCTGGTGGAGCCCATGGCAGCCGGATTCTCA	3735
QY	3952	TTGACGGCGTGGACATTTGCAGCATCGGCGCTGGAGGACTTGGCGGTCCAAGCTCTCAGTGA	4011
Db	3736	TTGACGGCGTGGACATTTGCAGCATCGGCGCTGGAGGACTTGGCGGTCCAAGCTCTCAGTGA	3795
QY	4012	TC C C T C A A G A T C C A G T G C T C T C A G G A A C C A T C A G A T T C A A C C T T T T G A C C	4071
Db	3796	T C C C T C A A G A T C C A G T G C T C T C A G G A A C C A T C A G A T T C A A C C T T A G C C	3855
QY	4072	G T C A C A C T G A C C A G A G A T C T G G G A T G C C T T G G A G A G G A C A T T C C T G A C C A A G G C C A T C T	4131
Db	3856	G T C A C A C T G A C C A G A G A T C T G G G A T G C C T T G G A G A G G A C A T T C C T G A C C A A G G C C A T C T	3915
QY	4132	C A A A G T T C C C C A A A A A G C T G C A T A C A G A T G T G G T G G A A A A C G G T G G A A A C T T C T C T G T G G	4191
Db	3916	C A A A G T T C C C C A A A A A G C T G C A T A C A G A T G T G G T G G A A A A C G G T G G A A A C T T C T C T G T G G	3975
QY	4192	G G G A G A G G C A G C T G C T C T G C A T T G C C A G S G S T G T G C T T C G C A A C T C C A A G A T C A T C C T T A	4251
Db	3976	G G G A G A G G C A G C T G C T C T G C A T T G C C A G S G S T G T G C T T C G C A A C T C C A A G A T C A T C C T T A	4035
QY	4252	T C G A T G A A G C C C A G A C C C T C C A T T G A C A T G G A C A G A C A C A C C C T G A T C C A G C G C A C A A T C C	4311
Db	4036	T C G A T G A A G C C C A G A C C C T C C A T T G A C A T G G A C A G A C A C A C C C T G A T C C A G C G C A C A A T C C	4095
QY	4312	G T G A A G C C T T C A G G G C T G C A C C G T G C T C G T C A T T G C C C A C C G T G T C A C C A C T G T G C T G A	4371
Db	4096	G T G A A G C C T T C A G G G C T G C A C C G T G C T C G T C A T T G C C C A C C G T G T C A C C A C T G T G C T G A	4155
QY	4372	A C T G T G A C C A C A T C C T G G T T A T G G G C A A T G G G A A G G T G G T A G A A T T G A T C G G C C G G A G G	4431
Db	4156	A C T G T G A C C A C A T C C T G G T T A T G G G C A A T G G G A A G G T G G T A G A A T T T G A T C G G C C G G A G G	4215
QY	4432	T A C T G C G G A A G A A G C C T G G G T C A T T G T T C G A G C C C C T C A T G G C C A C A G C C A C T T C T T C A C	4491
Db	4216	T A C T G C G G A A G A A G C C T G G G T C A T T G T T C G A G C C C C T C A T G G C C A C A G C C A C T T C T T C A C	4275
QY	4492	T G A G A T A A G G A G A T G T G G A G A C T T C A T G G A G G C T G G C A G C T G A G C T C A G A G G T T C A C A C A	4551
Db	4276	T G A G A T A A G G A G A T G T G G A G A C T T C A T G G A G G C T G G C A G C T G A G C T C A G A G G T T C A C A C A	4335
QY	4552	G G T G C A G C T T C A G G C C C C A C A G T C T G C G A C C T T C T T G T T T G G A G A T G A G A A C T T C T C C T G	4611
Db	4336	G G T G C A G C T T C A G G C C C C A C A G T C T G C G A C C T T C T T G T T T G G A G A T G A G A A C T T C T C C T G	4395
QY	4612	G A A G C A G G G T A A A T G T A G G G G G G T G G G A T T G C T G G A T G G A A A C C C T G G A A T A G G C T A	4671
Db	4396	G A A G C A G G G T A A A T G T A G G G G G G T G G G A T T G C T G G A T G G A A A C C C T G G A A T A G G C T A	4455
QY	4672	C T T G A T G G C T C T C A A G A C C T T A G A A C C C A G A A C C A T C T A A G A C A T G G G A T T C A G T G A T C	4731
Db	4456	C T T G A T G G C T C T C A A G A C C T T A G A A C C C A G A A C C A T C T A A G A C A T G G G A T T C A G T G A T C	4515
QY	4732	A T G T G T T C T C C T T T T A A C T T A C A T G C T G A A T A A T T T T A T A A T A A G G T A A A A G C T T A T A G	4791
Db	4516	A T G T G T T C T C C T T T T A A C T T A C A T G C T G A A T A A T T T T A T A A T A A G S T A A A A G C T T A T A G	4575
QY	4792	T T T T C T G A T C T G T T A G A A G T G T T G C A A A T G C T G T A C T G A C T T T G T A A A A A T A T A A A A C T	4851
Db	4576	T T T T C T G A T C T G T T A G A A G T G T T G C A A A T G C T G T A C T G A C T T T G T A A A A A T A T A A A A C T	4635
QY	4852	AAG 4854	
Db	4636	AAG 4638	

ABK92211	
ID	ABK92211 standard; DNA; 4427 BP.
XX	
AC	ABK92211;
XX	
DT	15-AUG-2002 (first entry)
XX	
DE	Prostate cancer-associated DNA sequence #97.
XX	
KW	Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
KW	gene therapy; gene; ds.
XX	
OS	Mammalia.
XX	
PN	WO200230268-A2.
XX	
PD	18-APR-2002.
XX	
PF	12-OCT-2001; 2001WO-US32045.
XX	
PR	13-OCT-2000; 2000US-0687576.
PR	08-DEC-2000; 2000US-0733288.
PR	08-DEC-2000; 2000US-0733742.
PR	24-JAN-2001; 2001US-263957P.
PR	16-MAR-2001; 2001US-276791P.
PR	16-MAR-2001; 2001US-276888P.
PR	06-APR-2001; 2001US-281922P.
PR	24-APR-2001; 2001US-286214P.
PR	30-APR-2001; 2001US-0847046.
PR	04-MAY-2001; 2001US-288589P.
XX	
PA	(EOSB-) EOS BIOTECHNOLOGY INC.
XX	
PI	Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX	
DR	WPI; 2002-471335/50.
DR	P-PSDB; ABG61895.
XX	
PT	Detecting a prostate cancer-associated transcript in a cell in a
PT	patient, useful for diagnosing prostate cancer (PC) or screening
PT	modulators of PC, by determining if prostate cancer-associated genes
PT	are expressed in a prostate tissue
XX	
PS	Claim 22; Page 380-381; 436pp; English.
XX	
CC	The present invention relates to methods of detecting a prostate
CC	cancer-associated transcript in a cell from a patient. The method
CC	comprises contacting a biological sample from the patient with
CC	prostate cancer-associated polynucleotides (designated PC genes) that
CC	selectively hybridise to a sequence that is at least 80% identical
CC	to them. The prostate cancer-associated polynucleotide sequences
CC	are differentially expressed in prostate tumour tissue or in
CC	prostate cancer and are derived from the tissues of various
CC	organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC	The methods of the invention are useful for diagnosing and treating
CC	prostate cancer in mammals. The prostate cancer-associated genes are
CC	useful for diagnosing or treating prostate cancer, as well as for
CC	identifying modulators of prostate cancer or agents that inhibit
CC	prostate cancer. The nucleic acid sequences are particularly useful
CC	in gene therapy, as a vaccine or in antisense applications.
CC	ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC	sequences.
XX	
SQ	Sequence 4427 BP; 1068 A; 1116 C; 1164 G; 1078 T; 1 other;

Query Match	87.4%;	Score 4249.6;	DB 24;	Length 4427;
Best Local Similarity	97.2%;	Pred. NO. 0;		
Matches 4386;	Conservative 1;	Mismatches 20;	Indels 105;	Gaps 2;
QY	351	ATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	410	
Db	1	ATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	60	

QY	411	ATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT	470
Db	61	ATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT	120
QY	471	GGCCCTCGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGAGGCGAGCTGTCCCACCG	530
Db	121	GGCCCTCGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGAGGCGAGCTGTCCCACCG	180
QY	531	TGGGGAGATATGATGCTGCCTTGAGAACCATGATTCCTTCCGTCCTCCAAAGCCGAGGTTT	590
Db	181	TGGGGAGATATGATGCTGCCTTGAGAACCATGATTCCTTCCGTCCTCCAAAGCCGAGGTTT	240
QY	591	CCTGCCCCAGCCCTGGACAATGCTGGCCCTGTTCTCCTACCTACCCGCTGTCATGGCTC	650
Db	241	CCTGCCCCAGCCCTGGACAATGCTGGCCCTGTTCTCCTACCTACCCGCTGTCATGGCTC	300
QY	651	ACCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACCATCCCTCCACTG	710
Db	301	ACCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACCATCCCTCCACTG	360
QY	711	TCAGTCCATGATGCCTCAGACAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA	770
Db	361	TCAGTCCATGATGCCTCAGACAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA	420
QY	771	GTCTAAGCGGAGGGATTGAAAAAGCTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGA	830
Db	421	GTCTAAGCGGAGGGATTGAAAAAGCTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGA	480
QY	831	ACAAGTTGATTTTCGATGCACTTCTGGGCATCTGCTTCTGCATFGCCAGTGTACTCGGG	890
Db	481	ACAAGTTGATTTTCGATGCACTTCTGGGCATCTGCTTCTGCATFGCCAGTGTACTCGGG	540
QY	891	CCAATATTGATATACCAAAGATCCTGGAATATTAGAAAGAGCAGTTGGGAATGTTGTC	950
Db	541	CCAATATTGATATACCAAAGATCCTGGAATATTAGAAAGAGCAGTTGGGAATGTTGTC	600
QY	951	CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGGAAGTCTCTGAGTTTC	1010
Db	601	CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGGAAGTCTCTGAGTTTC	660
QY	1011	TCCTCCAGTTGGATCATEAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCCTCC	1070
Db	661	TCCTCCAGTTGGATCATEAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCCTCC	720
QY	1071	TTTGGCCTTGAGAAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC	1130
Db	721	TTTGGCCTTGAGAAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC	780
QY	1131	ATCAGTCTCTTACCCGGTGATGATAAACTACCTGTTTGAAGGGTGCTATGGACCCCTA	1190
Db	781	ATCAGTCTCTTACCCGGTGATGATAAACTACCTGTTTGAAGGGTGCTATGGACCCCTA	840
QY	1191	GTACTGATCACCTGCGCATCGCTGCTCATCTGCAGCATTTCTTCCCTACTTCATTTATGGA	1250
Db	841	GTACTGATCACCTGCGCATCGCTGCTCATCTGCAGCATTTCTTCCCTACTTCATTTATGGA	900
QY	1251	TACACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTTCCTCCACTGGCGGTATTTCATG	1310
Db	901	TACACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTTCCTCCACTGGCGGTATTTCATG	960
QY	1311	ACAAGATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGAGGTCAGCGACCGCATCCGT	1370
Db	961	ACAAGATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGAGGTCAGCGACCGCATCCGT	1020
QY	1371	GTGACCAAGTGAAGTTCTCATTGCAATTAAGCTGATTAAAAATGTAACATGGGAGAAACCA	1430
Db	1021	GTGACCAAGTGAAGTTCTCATTGCAATTAAGCTGATTAAAAATGTAACATGGGAGAAACCA	1080
QY	1431	TTTGCAAAAAATCATTTGAAGACCTTAAGAAGGAAGGAAAGCAACTATTGGAGAAGTGCGGG	1490
Db	1081	TTTGCAAAAAATCATTTGAAG-----	1099
QY	1491	CTTGTCCAGAGCCTTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGGTC	1550

Db	1100	-----GTATGGAAAGTCTGACTTTCGTCTCCAAACCTGGTGATGGC-----	1140
QY	1551	TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG	1610
Db	1141	-----ATGGCCTTCAGCATG	1155
QY	1611	CTGGCCTCCTTGAATCTCCTTCGGCTGTTCAGTGTTCTTTGTGCCCTATTGCAGTCAAAGGT	1670
Db	1156	CTGGCCTCCTTGAATCTCCTTCGGCTGTTCAGTGTTCTTTGTGCCCTATTGCAGTCAAAGGT	1215
QY	1671	CTCACGAATTCCAAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTCCTCCAGGAGACCCCT	1730
Db	1216	CTCACGAATTCCAAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTCCTCCAGGAGACCCCT	1275
QY	1731	GTTTCTCATGTCCAGACATTAACAAGACCCCCAGCAAAAGCTCTGGTCTTTGAGGAGGCCACC	1790
Db	1276	GTTTCTCATGTCCAGACATTAACAAGACCCCCAGCAAAAGCTCTGGTCTTTGAGGAGGCCACC	1335
QY	1791	TTGTCTATGGCAACAGACCTGTCCCCGGGATCGTCAATGGGGGCACTCGAGCTGGAGAGGAAC	1850
Db	1336	TTGTCTATGGCAACAGACCTGTCCCCGGGATCGTCAATGGGGGCACTCGAGCTGGAGAGGAAC	1395
QY	1851	GGGCATGCTTCTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGCCACAGGAAGAAGGG	1910
Db	1396	GGGCATGCTTCTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGCCACAGGAAGAAGGG	1455
QY	1911	AACAGCCTGGGCCACAGAGTTGCACAAAGATCAACCTGGTGGTGCTCAAGGGGATGATGTTA	1970
Db	1456	AACAGCCTGGGCCACAGAGTTGCACAAAGATCAACCTGGTGGTGCTCAAGGGGATGATGTTA	1515
QY	1971	GGGTCCTCGCGCAACACAGGGAGTGGTAAGAGCAGCCCTGTTGTACGCCATCCTGGAGGAG	2030
Db	1516	GGGTCCTCGCGCAACACAGGGAGTGGTAAGAGCAGCCCTGTTGTACGCCATCCTGGAGGAG	1575
QY	2031	ATGCACCTGCTCGAGGGCTCGGTGGGGTGAGGAAAGCCCTGGCCTATGTCCCCCAGCAG	2090
Db	1576	ATGCACCTGCTCGAGGGCTCGGTGGGGTGAGGAAAGCCCTGGCCTATGTCCCCCAGCAG	1635
QY	2091	GCCTGGATCGTCAGCGGGAACATCAGGGAGAACATCCTCATGGAGGCGCATATGACAAG	2150
Db	1636	GCCTGGATCGTCAGCGGGAACATCAGGGAGAACATCCTCATGGAGGCGCATATGACAAG	1695
QY	2151	GCCCGATACCTCCAGGTGCTCCACTGCTCCTCCCTGAATCGGGACCTTGTGCCC	2210
Db	1696	GCCCGATACCTCCAGGTGCTCCACTGCTCCTCCCTGAATCGGGACCTTGTGCCC	1755
QY	2211	TTTGGACACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGCGAGAAACAG	2270
Db	1756	TTTGGAGACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGCGAGAAACAG	1815
QY	2271	AGGATCAGCCTGGCCCGCGCGTCTATTCCGACCCGTGAGATCTACCTGCTGGACGACCCC	2330
Db	1816	AGGATCAGCCTGGCCCGCGCGTCTATTCCGACCCGTGAGATCTACCTGCTGGACGACCCC	1875
QY	2331	CTGTCTGTGGACGCGCCACGTGGGGAAGCACATTTTGGAGAGTGCATTAAGAAGACA	2390
Db	1876	CTGTCTGTGGACGCGCCACGTGGGGAAGCACATTTTGGAGAGTGCATTAAGAAGACA	1935
QY	2391	CTCAGGGGAAGACGGTCTGCTGGTGACCCACCCAGCTGCAGTACTTAGAATTTGTGGC	2450
Db	1936	CTCAGGGGAAGACGGTCTGCTGGTGACCCACCCAGCTGCAGTACTTAGAATTTGTGGC	1995
QY	2451	CAGATCATTTTGTGGAAAAATGGGAAAAATCTGTGAAAAATGGAACCTCACAGTGAGTTAATG	2510
Db	1996	CAGATCATTTTGTGGAAAAATGGGAAAAATCTGTGAAAAATGGAACCTCACAGTGAGTTAATG	2055
QY	2511	CAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAGCCACTTCGGAC	2570
Db	2056	CAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAGCCACTTCGGAC	2115
QY	2571	ATGTTGCAGGACACAGCAAAAGATAGCAGAGAAGCCAAAGGTAGAAAAGTCAAGGCTCGGCC	2630

Db 2116 ATGTTGCAGGACACAGCAAAGATAGCAGAGAGCCAAAGGTAGAAAAGTCAGGGCTCTGGCC 2175

QY 2631 ACCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCCGGAGCATCAGCTCACACAGGAG 2690

Db 2176 ACCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCCGGAGCATCAGCTCACACAGGAG 2235

QY 2691 GAGGAGATGGAAGAGGCTCCTTGAGTTGGAGGCTCTACCACCACCTACATCCAGGCAGCT 2750

Db 2236 GAGGAGATGGAAGAGGCTCCTTGAGTTGGAGGCTCTACCACCACCTACATCCAGGCAGCT 2295

QY 2751 GGAGGTTACATGGTCTCTTGCATAAATTTTCTTCTCGTGGTGCTGATCGTCTCTTAACG 2810

Db 2296 GGAGGTTACATGGTCTCTTGCATAAATTTTCTTCTCGTGGTGCTGATCGTCTCTTAACG 2355

QY 2811 ATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGCAGC 2870

Db 2356 ATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGCAGC 2415

QY 2871 CGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCGAGACAATCCTCAACTGTCC 2930

Db 2416 CGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCGAGACAATCCTCAACTGTCC 2475

QY 2931 TTCTACCAGCTGGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCC 2990

Db 2476 TTCTACCAGCTGGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCC 2535

QY 2991 TCAGGGATTTTCACAAAGTCACGAGGAGGCAATCCACGGCCCTGCACACAAGCTCTTC 3050

Db 2536 TCAGGGATTTTCACAAAGTCACGAGGAGGCAATCCACGGCCCTGCACACAAGCTCTTC 2595

QY 3051 AACAAAGTTTTCCGCTGCCCATGAGTTTTCTTGACACCATCCCAATAGGCCGGCTTTTG 3110

Db 2596 AACAAAGTTTTCCGCTGCCCATGAGTTTTCTTGACACCATCCCAATAGGCCGGCTTTTG 2655

QY 3111 AACTGCTTCGCAGGGGACTTGGAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAGCAG 3170

Db 2656 AACTGCTTCGCAGGGGACTTGGAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAGCAG 2715

QY 3171 TTCCTGGTCCTTAATGGTGATCGCCGTCCTGTTGATTGTCAGTGCTGTCTCCA 3230

Db 2716 TTCCTGGTCCTTAATGGTGATCGCCGTCCTGTTGATTGTCAGTGCTGTCTCCA 2775

QY 3231 TATATCCTGTTAATGGGAGCCATAATCATGGTTATTTGCTTCAATTTATATATGATGTTT 3290

Db 2776 TATATCCTGTTAATGGGAGCCATAATCATGGTTATTTGCTTCAATTTATATATGATGTTT 2835

QY 3291 AAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTATTTCTCC 3350

Db 2836 AAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTATTTCTCC 2895

QY 3351 CACATCCTCAATTTCTCTGCAAGGCCCTGAGCTCCATCCATGCTCTATGGAAAAACTGAAGAC 3410

Db 2896 CACATCCTCAATTTCTCTGCAAGGCCCTGAGCTCCATCCATGCTCTATGGAAAAACTGAAGAC 2955

QY 3411 TTCAATCAGCCAGTTAAGAGGCTGACTGATGCGCAGAAATAACTACCTGCTGTTGTTTCTA 3470

Db 2956 TTCAATCAGCCAGTTAAGAGGCTGACTGATGCGCAGAAATAACTACCTGCTGTTGTTTCTA 3015

QY 3471 TCTTCCACACGATGGATGGCATTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT 3530

Db 3016 TCTTCCACACGATGGATGGCATTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT 3075

QY 3531 GTTGCCCTGTTCCGTGGCTTTTGGCAATTTCTCCACCCCTACTCCTTTAAAGTCATGGCT 3590

Db 3076 GTTGCCCTGTTCCGTGGCTTTTGGCAATTTCTCCACCCCTACTCCTTTAAAGTCATGGCT 3135

QY 3591 GTCAACATCGTGTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCCCCGATTTGGGAG 3650

Db 3136 GTCAACATCGTGTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCCCCGATTTGGCTTGGAG 3195

QY 3651 ACAGAGGCACAGTTCACGGCTGTAGAGAGGATACTGCAGTACATGAAGATGTGTCTCG 3710

Db 3196 ACAGAGGCACAGTTCACGGCTGTAGAGAGGATACTGCAGTACATGAAGATGTGTCTCG 3255

QY 3711 GAAGCTCCTTTACACATGGAAGGCACAAAGTTGTCCCCAGGGGTGGCCACAGCATTCGGGAA 3770

Db 3256 GAAGCTCCTTTACACATGGAAGGCACAAAGTTGTCCCCAGGGGTGGCCACAGCATTCGGGAA 3315

QY 3771 ATCATATTTTCAGGATTCACATGAAATACAGAGACAACACACCCACCGTCTTCACGGC 3830

Db 3316 ATCATATTTTCAGGATTCACATGAAATACAGAGACAACACACCCACCGTCTTCACGGC 3375

QY 3831 ATCAACCTGACCAATCCGCGGCCACGAAGTGGTGGGCATCGTGGGAAGGACGGGCTTGGG 3890

Db 3376 ATCAACCTGACCAATCCGCGGCCACGAAGTGGTGGGCATCGTGGGAAGGACGGGCTTGGG 3435

QY 3891 AAGTCCCTCCTTGGGCATGGCTCTCTTCCGCCCTGGTGGAGCCCATGGCAGGCCGGATTTCTC 3950

Db 3436 AAGTCCCTCCTTGGGCATGGCTCTCTTCCGCCCTGGTGGAGCCCATGGCAGGCCGGATTTCTC 3495

QY 3951 ATTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAGGACTTGCGGTCCAAGCTCTCAGTG 4010

Db 3496 ATTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAGGACTTGCGGTCCAAGCTCTCAGTG 3555

QY 4011 ATCCCTCAAGATCCAGTGCTGCTCTCAGGAACCATCAGATTCAAACCTAGATCCCCTTTGAC 4070

Db 3556 ATCCCTCAAGATCCAGTGCTGCTCTCAGGAACCATCAGATTCAAACCTAGATCCCCTTTGAC 3615

QY 4071 CGTCACACTGACCAGAGATCTGGGATGCCCTTGAGAGGACATTCCTGACCAAGGCCCATC 4130

Db 3616 CGTCACACTGACCAGAGATCTGGGATGCCCTTGAGAGGACATTCCTGACCAAGGCCCATC 3675

QY 4131 TCAAAGTTCCCCAAAAAGCTGCATACAGATGTGGTGGAAAAACGGTGGAAACTTCTCTGTG 4190

Db 3676 TCAAAGTTCCCCAAAAAGCTGCATACAGATGTGGTGGAAAAACGGTGGAAACTTCTCTGTG 3735

QY 4191 GGGGAGAGGCAGCTGCTCTGCATTTGCCAGGGGCTGTGCTTCGCAACTCCAAGATCATCCTT 4250

Db 3736 GGGGAGAGGCAGCTGCTCTGCATTTGCCAGGGGCTGTGCTTCGCAACTCCAAGATCATCCTT 3795

QY 4251 ATCGATGAAGCCACAGCCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATC 4310

Db 3796 ATCGATGAAGCCACAGCCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATC 3855

QY 4311 CGTGAAGCCCTTCAGGGGCTGCACCGTGTCTCGTCAATTGCCCAACCGTGTCACCACTGTGCTG 4370

Db 3856 CGTGAAGCCCTTCAGGGGCTGCACCGTGTCTCGTCAATTGCCCAACCGTGTCACCACTGTGCTG 3915

QY 4371 AACTGTGACCCACATCCTTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGCCCGGAG 4430

Db 3916 AACTGTGACCCACATCCTTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGCCCGGAG 3975

QY 4431 GTACTGGGGAAGAACCTTGGGTCAATTTGTCGAGCCCTCATGGCCACAGCCACTTCTTCA 4490

Db 3976 GTACTGGGGAAGAACCTTGGGTCAATTTGTCGAGCCCTCATGGCCACAGCCACTTCTTCA 4035

QY 4491 CTGAGATAAGGAGATGTGGAGACTTCATGGAGGCTGGCAGCTGAGCTCAGAGTTCAACAC 4550

Db 4036 CTGAGATAAGGAGATGTGGAGACTTCATGGAGGCTGGCAGCTGAGCTCAGAGTTCAACAC 4095

QY 4551 AGGTGCAGCTTCGAGGCCCCACAGTCTGCGACCTTCTTGTTTGGAGATGAGAACTTCTCCT 4610

Db 4096 AGGTGCAGCTTCGAGGCCCCACAGTCTGCGACCTTCTTGTTTGGAGATGAGAACTTCTCCT 4155

QY 4611 GGAAGCAGGGGTAAATGTAGGGGGGTGGGGATTTGCTGGATGGAAACCCCTGGAATAGGCT 4670

Db 4156 GGAAGCAGGGGTAAATGTAGGGGGGTGGGGATTTGCTGGATGGAAACCCCTGGAATAGGCT 4215

QY 4671 ACTTGATGGCTCTCAAGACCTTAGAACCCCGAGAACCCATCTAAGACATGGATTCAGTGAT 4730

Db 4216 ACTTGATGGCTCTCAAGACCTTAGAACCCCGAGAACCCATCTAAGACATGGATTCAGTGAT 4275

QY 4731 CATGTGGTCTCCTTTTAACTTACATGCTGAATAAATTTTATAATAAGGTAAAAGCTTATA 4790

Db 4276 CATGTGGTCTCCTTTTAACTTACATGCTGAATAAATTTTATAATAAGGTAAAAGCTTATA 4335

QY 4791 GTTTCTGATCTGTGTAGAAAGTGTGCAAAATGCTGTACTGACTTTGTAAAAATATAAAC 4850
|||||
Db 4336 GTTTCTGATCTGTGTTAGAAGTGTGCAAAATGCTGTACTGACTTTGTAAAAATATAAAC 4395

QY 4851 TAAGGAAAAC TC 4862
|||||
Db 4396 TAAGGAAAAC TC 4407

RESULT 3
AAF83643
ID AAF83643 standard; cDNA; 4149 BP.
XX
AC AAF83643;
XX
DT 23-JUL-2001 (first entry)
XX
DE Novel human transporter protein (NHP) encoding cDNA.
XX
KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX

FH Key Location/Qualifiers
FT CDS 1..4149
FT /*tag= a
FT misc_feature 259..267
FT /*tag= b
FT /note= "the amino acid residues encoded by the above 3
FT codons are not indicated in the corresponding
FT protein sequence"
XX
PN WO200132706-A2.
XX
PD 10-MAY-2001.

XX
PF 31-OCT-2000; 2000WO-US29852.
XX
PR 02-NOV-1999; 99US-0163018.

XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
PI
XX
DR WPI; 2001-343477/36.
DR P-PSDB; AAB62555.
XX

PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
PT
XX
PS Disclosure; Page 44-46; 59pp; English.
XX

CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAF83635-647 represent NHP nucleic acid sequences.
XX
SQ Sequence 4149 BP; 982 A; 1080 C; 1091 G; 996 T; 0 other;

Query Match 85.2%; Score 4142.6; DB 22; Length 4149;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4145; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 351 ATGACTAGGAAGAGGACATACTGGGTGCCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 410
|||||
Db 1 ATGACTAGGAAGAGGACATACTGGGTGCCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 50

QY 411 ATCGACATAGCGGATGACATGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT 470
|||||
Db 61 ATCGACATAGCGGATGACATGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT 120

QY 471 GGCCCTGGAGTCAGCAAGAGAGAAAATCCTGAGGCTCCAGGAGGGCAGCTGTCCCACCG 530
|||||
Db 121 GGCCCTGGAGTCAGCAAGAGAGAAAATCCTGAGGCTCCAGGAGGGCAGCTGTCCCACCG 180

QY 531 TGGGGGAAGTATGATGCTGCCTTGAGAAACCATGATTCCCTTCCGTCCCAAGCCGAGGTTT 590
|||||
Db 181 TGGGGGAAGTATGATGCTGCCTTGAGAAACCATGATTCCCTTCCGTCCCAAGCCGAGGTTT 240

QY 591 CCTGCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTCTACCTCACCGTGTCATGGCTC 550
|||||
Db 241 CCTGCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTCTACCTCACCGTGTCATGGCTC 300

QY 651 ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAAACACCATCCCTCCACTG 710
|||||
Db 301 ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAAACACCATCCCTCCACTG 360

QY 711 TCAGTCCATGATGCCTCAGACAAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA 770
|||||
Db 361 TCAGTCCATGATGCCTCAGACAAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA 420

QY 771 GTCTCAAGGCGAGGGATTGAAAAGCTTTCAGTGCCTTCTGGTGATGCTGAGGTTCCAGAGA 830
|||||
Db 421 GTCTCAAGGCGAGGGATTGAAAAGCTTTCAGTGCCTTCTGGTGATGCTGAGGTTCCAGAGA 480

QY 831 ACAAGGTTGATTTTCGATGCACCTCTGGGGCATCTGCTTCTCCATTCGCCAGTGACTCGGG 890
|||||
Db 481 ACAAGGTTGATTTTCGATGCACCTCTGGGGCATCTGCTTCTCCATTCGCCAGTGACTCGGG 540

QY 891 CCAATATTGATTATACCAAAGATCCTGGAAATATTCAGAGAGAGCAGTTGGGGAATGTTGTC 950
|||||
Db 541 CCAATATTGATTATACCAAAGATCCTGGAAATATTCAGAGAGAGCAGTTGGGGAATGTTGTC 600

QY 951 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC 1010
|||||
Db 601 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC 560

QY 1011 TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTCCCTCC 1070
|||||
Db 661 TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCCAAGCAGCTGTTCCCTCC 720

QY 1071 TTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC 1130
|||||
Db 721 TTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC 780

QY 1131 ATCAGCTTCTTCACCCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTA 1190
|||||
Db 781 ATCAGCTTCTTCACCCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTA 840

QY 1191 GTACTGATCACCTGGGCATCGCTGGTCATCTGCAGCATTTCTTCCTACTTCAATTATGGA 1250
|||||
Db 841 GTACTGATCACCTGGGCATCGCTGGTCATCTGCAGCATTTCTTCCTACTTCAATTATGGA 900

QY 1251 TACACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCAGTGGCGGTATTTCATG 1310
|||||
Db 901 TACACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCAGTGGAGGTATTTCATG 960

QY 1311 ACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT 1370
|||||
Db 961 ACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT 1020

QY 1371 GTGACCAGTGAAGTTCTCACTTGCATTTAAGCTGATTAAATGTACACATGGGAGAAACCA 1430
|||||
Db 1021 GTGACCAGTGAAGTTCTCACTTGCATTTAAGCTGATTAAATGTACACATGGGAGAAACCA 1080

QY 1431 TTTGCAAAAATCATTGAACACCTAAGAAAGGAAGAAACACTATTGGAGAAAGTCGGGG 1490

QY 3651 ACAGAGGCACAGTTTCACGGCTGTAGAGAGGATACTGCAGTACATGAAGATGTGTGTCG 3710
|||||
Db 3301 ACAGAGGCACAGTTCACGGCTGTAGAGAGGATACTGCAGTACATGAAGATGTGTGTCG 3360

QY 3711 GAAGCTCCTTTACACATGGAAGGCACAAAGTTGTCCCCAGGGGTGGCCACAGCATGGGAA 3770
|||||
Db 3361 GAAGCTCCTTTACACATGGAAGGCACAAAGTTGTCCCCAGGGGTGGCCACAGCATGGGAA 3420

QY 3771 ATCATATTTCAGGATTATCACATGAAATACAGAGACAACACACCCACCCTGCTTCACGGC 3830
|||||
Db 3421 ATCATATTTCAGGATTATCACATGAAATACAGAGACAACACACCCACCCTGCTTCACGGC 3480

QY 3831 ATCAACCTGACCATCCCGCGGCCACGAAGTGGTGGGCATCTGGGAAGGACGGGCTCTGG 3890
|||||
Db 3481 ATCAACCTGACCATCCCGCGGCCACGAAGTGGTGGGCATCTGGGAAGGACGGGCTCTGG 3540

QY 3891 AAGTCCTCTCTTGGGCATGGCTCTCTCCGCCCTGGTGGAGCCCATGGCAGGCCGGATTCTC 3950
|||||
Db 3541 AAGTCCTCTCTTGGGCATGGCTCTCTCCGCCCTGGTGGAGCCCATGGCAGGCCGGATTCTC 3600

QY 3951 ATTGACGGCGTGGACATTTGCAGCATCGGCCTGGAGGACTTGGCGTCCAAGCTCTCAGTG 4010
|||||
Db 3601 ATTGACGGCGTGGACATTTGCAGCATCGGCCTGGAGGACTTGGCGTCCAAGCTCTCAGTG 3660

QY 4011 ATCCCTCAAGATCCAGTGTCTCTCAGGAACCATCAGATTCAACCTTAGATCCCTTTGAC 4070
|||||
Db 3661 ATCCCTCAAGATCCAGTGTCTCTCAGGAACCATCAGATTCAACCTTAGATCCCTTTGAC 3720

QY 4071 CGTCACACTGACCAGCAGATCTGGGATGCCTTGGAGAGGACATTCCTGACCAAGGCCATC 4130
|||||
Db 3721 CGTCACACTGACCAGCAGATCTGGGATGCCTTGGAGAGGACATTCCTGACCAAGGCCATC 3780

QY 4131 TCAAAGTTCCTCCCAAAAAGCTGCATACAGATGTGTGGAAACGGTGGAACCTTCTCTGTG 4190
|||||
Db 3781 TCAAAGTTCCTCCCAAAAAGCTGCATACAGATGTGTGGAAACGGTGGAACCTTCTCTGTG 3840

QY 4191 GGGAGAGGCAGCTGCTCTGCAATTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCCTT 4250
|||||
Db 3841 GGGAGAGGCAGCTGCTCTGCAATTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCCTT 3900

QY 4251 ATCGATGAAGCCACAGCCTCCATTTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATC 4310
|||||
Db 3901 ATCGATGAAGCCACAGCCTCCATTTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATC 3960

QY 4311 CGTGAAGCCTTCCAGGGCTGCACCGTGTCTGTCATTGCCACCGTGTCAACCACTGTGCTG 4370
|||||
Db 3961 CGTGAAGCCTTCCAGGGCTGCACCGTGTCTGTCATTGCCACCGTGTCAACCACTGTGCTG 4020

QY 4371 AACTGTGACCACATCCTGGTTATGGGCAATGGGAAGGTGTTAGAAATTTGATCGGCCCGGAG 4430
|||||
Db 4021 AACTGTGACCACATCCTGGTTATGGGCAATGGGAAGGTGTTAGAAATTTGATCGGCCCGGAG 4080

QY 4431 GTACTCGGGAAGAACCTGGGTCAATTGTCGAGCCCTCATGGCCACAGCCACTTCTTCA 4490
|||||
Db 4081 GTACTCGGGAAGAACCTGGGTCAATTGTCGAGCCCTCATGGCCACAGCCACTTCTTCA 4140

QY 4491 CTGAGATAA 4499
|||||
Db 4141 CTGAGATAA 4149

RESULT 4
AAF83647
ID AAF83647 standard; cDNA; 5020 BP.
XX
AC AAF83647;

XX 23-JUL-2001 (first entry)

XX Novel human transporter protein (NHP) nucleic acid sequence.

XX Novel human protein; transporter protein; NHP; therapeutic; diagnostic;

KW gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200132706-A2.
XX
PD 10-MAY-2001.
XX
PF 31-OCT-2000; 2000WO-US29852.
XX
PR 02-NOV-1999; 99US-0163018.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI; 2001-343477/36.
XX
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT -
XX
PS Disclosure; Page 57-59; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAF83635-647 represent NHP nucleic acid sequences.
XX
SQ Sequence 5020 BP; 1240 A; 1249 C; 1329 G; 1202 T; 0 other;

Query Match 85.0%; Score 4133.6; DB 22; Length 5020;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 4529; Conservative 0; Mismatches 4; Indels 363; Gaps 3;

QY 330 AAGAATTGGCAGGAACTGAAATGACTAGGAAGAGGACATACTGGTGCCCCAACTCTTCT 389
|||||
Db 99 AAGAATTGGCAGGAACTGAAATGACTAGGAAGAGGACATACTGGTGCCCCAACTCTTCT 158

QY 390 GGTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTTAT 449
|||||
Db 159 GGTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTTAT 218

QY 450 AAAACCTATACTCTCCAAGATGGCCCCCTGGAGTCAGCAAGAGAGAAAATCCTGAGGCTCCA 509
|||||
Db 219 AAAACCTATACTCTCCAAGATGGCCCCCTGGAGTCAGCAAGAGAGAAAATCCTGAGGCTCCA 278

QY 510 GGGAGGGCAGCTGTCCCACCGTGGGGGAAGTATGATGCTGCTTGAGAACCATGATTTCCC 569
|||||
Db 279 GGGAGGGCAGCTGTCCCACCGTGGGGGAAGTATGATGCTGCTTGAGAACCATGATTTCCC 338

QY 570 TTCCGTCCCAAGCCGAGGTTTCTGCCCCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTCC 629
|||||
Db 339 TTCCGTCCCAAGCCGAGGTTTCTGCCCCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTCC 398

QY 630 TACCTCACCCGTGTCATGGCTCACCCCCGCTCATGATCCCAAAGCTTACGGAGTCGCTTAGAT 689
|||||
Db 399 TACCTCACCCGTGTCATGGCTCACCCCCGCTCATGATCCCAAAGCTTACGGAGTCGCTTAGAT 458

QY 690 GAGAACACCATCCCTCCACTGTCAGTCCCATGATGCCTCAGACAAAAAATGTCCAAAGGCTT 749
|||||
Db 459 GAGAACACCATCCCTCCACTGTCAGTCCCATGATGCCTCAGACAAAAAATGTCCAAAGGCTT 518

QY 750 CACCGCCTTTGGGAAGAAGACTCTCAAGCGGAGGGATTGAAAAAGCTTCAGTGTCTG 809
|||||

Db 519 CACCGCCTTTGGGAAGAAAGTCTCAAGGCGAGGGATTGAAAAAGCTTCAGTGTCTG 578

QY 810 GTGATGCTGAGGTTCCAGAGAACAAGGTGATTTTCGATGCACCTTCGGGCATCTGCTTC 869

Db 579 GTGATGCTGAGGTTCCAGAGAACAAGGTGATTTTCGATGCACCTTCGGGCATCTGCTTC 638

QY 870 TGCATTGCCAGTGTACTCGGGCCA----- 893

Db 639 TGCATTGCCAGTGTACTCGGGCCAGACTGCACCAATGGGGGAGGTGGGAGTCTTCCCAGA 698

QY 894 -----ATATTGATTATACCAAAGAT 913

Db 699 AGAAAAC TGAGGCACTGGTACCAGCTACGGGGAGGTGAAGATATTGATTATACCAAAGAT 758

QY 914 CCTGGAATATT CAGAAGAGCAGTTGGGGAATGTGTCCATGGAGTGGGACTCTGCTTTGC 973

Db 759 CCTGGAATATT CAGAAGAGCAGTTGGGGAATGTGTCCATGGAGTGGGACTCTGCTTTGC 818

QY 974 CCTTTTCTCCGGAATGTGTGAAGTCTCTGAGTTTCTCCTCCAGTGGATCATCAACCA 1033

Db 819 CCTTTTCTCCGGAATGTGTGAAGTCTCTGAGTTTCTCCTCCAGTGGATCATCAACCA 878

QY 1034 ACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCCTCCTTTGCTTTCAGAAAGCTCATCCA 1093

Db 879 ACGCACAGCCATCAGGTTCCAAGCAGCTGTTTCCCTCCTTTGCTTTCAGAAAGCTCATCCA 938

QY 1094 ATTTAAGTCTGTAATACACATCACCTCAGSAGAGGCCATCAGCTTCTTCACCGGTGATGT 1153

Db 939 ATTTAAGTCTGTAATACACATCACCTCAGSAGAGGCCATCAGCTTCTTCACCGGTGATGT 998

QY 1154 AAAC TACCTGTTGAAGGGGTGTGCTATGGACCCCTAGTACTGATCACCTGCGCATCGCT 1213

Db 999 AAAC TACCTGTTGAAGGGGTGTGCTATGGACCCCTAGTACTGATCACCTGCGCATCGCT 1058

QY 1214 GGT CATCTGCAGCATTTCTTCC TACTTCATTTATTTGGATACACTGCA TTTATTTGCCATCTT 1273

Db 1059 GGT CATCTGCAGCATTTCTTCC TACTTCATTTATTTGGATACACTGCA TTTATTTGCCATCTT 1118

QY 1274 ATGCTATCTCTCGGTTTTTCCCAC TGGCGGTATTCATGACAAGAA TGGCTGTGAAGGCTCA 1333

Db 1119 ATGCTATCTCTCGGTTTTTCCCAC TGGAGGTATTCATGACAAGAA TGGCTGTGAAGGCTCA 1178

QY 1334 GCATCACACATCTGAGGTCAGCGAC CAGCCATCCGTGTGACCAGTGAAGTTCTCACTTG 1393

Db 1179 GCATCACACATCTGAGGTCAGCGAC CAGCCATCCGTGTGACCAGTGAAGTTCTCACTTG 1238

QY 1394 CATTAAGCTGATTAANAATGTACACATGGGAGAAACCATTTTGCAAAAATCATTTGAAGACCT 1453

Db 1239 CATTAAGCTGATTAANAATGTACACATGGGAGAAACCATTTTGCAAAAATCATTTGAAGACCT 1298

QY 1454 AAGAAGGAAGGAAGAACTATTGGAGAAGTGGCGGCTTGTCAGAGCCCTGACAAGTAT 1513

Db 1299 AAGAAGGAAGGAAGAAAGCTATTGGAGAAGTGGCGGCTTGTCAGAGCCCTGACAAGTAT 1358

QY 1514 AACCTTGTTTCATCATCCCCACAGTGGCCACAGCGGTCTGGGTTCTCATCCACACATCCTT 1573

Db 1359 AACCTTGTTTCATCATCCCCACAGTGGCCACAGCGGTCTGGGTTCTCATCCACACATCCTT 1418

QY 1574 AAAGCTGAACACTCACAGCGTCAATGGCCTTCAGCATGCTGGCCTTCAGATCTCCTTCG 1633

Db 1419 AAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATGCTGGCCTTCAGATCTCCTTCG 1478

QY 1634 GCTGT CAGTGTCTTTGTGCC TATTGCAGTCAAAGGTCTCACGAATTCCAAAGTCTGCAGT 1693

Db 1479 GCTGT CAGTGTCTTTGTGCC TATTGCAGTCAAAGGTCTCACGAATTCCAAAGTCTGCAGT 1538

QY 1694 GATGAGGTTCAAGAAAGTTTTTCCCTCCAGSAGAGCCCTGTTTTCTATGTCAGACATTACA 1753

Db 1539 GATGAGGTTCAAGAAAGTTTTTCCCTCCAGSAGAGCCCTGTTTTCTATGTCAGACATTACA 1598

QY 1754 AGACCCCGCAAAAGCTCTGGTCTTTTGAGGAGGCCACCTTGT CATGSCAACAGACCTGTCC 1813

Db 1599 AGACCCCGCAAAAGCTCTGGTCTTTTGAGGAGGCCACCTTGT CATGSCAACAGACCTGTCC 1658

QY 1814 CGGGATCGTCAATGGGGCAC TGGAGCTGGAGAGGAACGGGCATGCTTCTGAGGGGATGAC 1873

Db 1659 CGGGATCGTCAATGGGGCAC TGGAGCTGGAGAGGAACGGGCATGCTTCTGAGGGGATGAC 1718

QY 1874 CAGGCTTAGAGATGCCCTCGGGCCAGAGAGAAGGGAACAGCCTGGGCCCAGAGTTGCA 1933

Db 1719 CAGGCTTAGAGATGCCCTCGGGCCAGAGAGAAGGGAACAGCCTGGGCCCAGAGTTGCA 1778

QY 1934 CAAGATCAACCTGGTGTGTCCAAAGGGGATGATTTAGGGGTCTCGGCCAACACGGGGAG 1993

Db 1779 CAAGATCAACCTGGTGTGTCCAAAGGGGATGATTTAGGGGTCTCGGCCAACACGGGGAG 1838

QY 1994 TGGTAAGAGCAGCCCTGTTGT CAGCCATCCTGGAGGAGATGCAC TTTGCTCGAGGGCTCGGT 2053

Db 1839 TGGTAAGAGCAGCCCTGTTGT CAGCCATCCTGGAGGAGATGCAC TTTGCTCGAGGGCTCGGT 1898

QY 2054 GGGGTGCAGGGAAGCCTGGCC TATGTCCCCCAGCAGGCCCTGGATCTCAGCGGGAACAT 2113

Db 1899 GGGGTGCAGGGAAGCCTGGCC TATGTCCCCCAGCAGGCCCTGGATCTCAGCGGGAACAT 1958

QY 2114 CAGGAGAAACATCCTCATGGG AGGCGCATATGACAAGGCCCGATACCTCCAGGTGCTCCA 2173

Db 1959 CAGGAGAAACATCCTCATGGG AGGCGCATATGACAAGGCCCGATACCTCCAGGTGCTCCA 2018

QY 2174 CTGCTGCTCCCTGAATCGGG ACCTGGAACTTCTGCCCTTTGGAGACATGACAGAGATTGG 2233

Db 2019 CTGCTGCTCCCTGAATCGGG ACCTGGAACTTCTGCCCTTTGGAGACATGACAGAGATTGG 2078

QY 2234 AGAGCGGGCCTCAACCTCTCTG GGGGCGAGAAACAGAGGATCAGCCTGGCCCGCGCCGT 2293

Db 2079 AGAGCGGGCCTCAACCTCTCTG GGGGCGAGAAACAGAGGATCAGCCTGGCCCGCGCCGT 2138

QY 2294 CTATTCCGACCGTCAGATCTAC CTGCTGCAGCACCCCTGTCTGCTGCGACGCCACGT 2353

Db 2139 CTATTCCGACCGTCAGATCTAC CTGCTGCAGCACCCCTGTCTGCTGCGACGCCACGT 2198

QY 2354 GGGGAAGCACATTTT TGAGGAGTGCATTAAGAAGACACTCAGGGGGAAGACGGTCGTCCCT 2413

Db 2199 GGGGAAGCACATTTT TGAGGAGTGCATTAAGAAGACACTCAGGGGGAAGACGGTCGTCCCT 2258

QY 2414 GGTGACCCACAGCTGCAGTAC TTTAGAAATTTTGTGGCCAGATCATTTTGTGAAAAATGG 2473

Db 2259 GGTGACCCACAGCTGCAGTAC TTTAGAAATTTTGTGGCCAGATCATTTTGTGAAAAATGG 2318

QY 2474 GAAAACTGTGAAAAATGGAACT CACAGTCAGTTAATGCAGAAAAAGGGGAATATGCCCA 2533

Db 2319 GAAAACTGTGAAAAATGGAACT CACAGTCAGTTAATGCAGAAAAAGGGGAATATGCCCA 2378

QY 2534 ACTTATCCAGAAAGATGCACA AGGCCACTTCGGACATGTTGCAGGACACACAGCAAAAGAT 2593

Db 2379 ACTTATCCAGAAAGATGCACA AGGCCACTTCGGACATGTTGCAGGACACACAGCAAAAGAT 2438

QY 2594 AGCAGAGAACCCAAAGGTAGAA AGTCAAGGCTCTGGCCACCTCCCTGGAAAGAGTCTCTCAA 2653

Db 2439 AGCAGAGAACCCAAAGGTAGAA AGTCAAGGCTCTGGCCACCTCCCTGGAAAGAGTCTCTCAA 2498

QY 2654 CGGAAATGCTGTGCCGGAGCAT CAGCTCACACAGGAGGAGAGATGGAAGAAGGCTCCTT 2713

Db 2499 CGGAAATGCTGTGCCGGAGCAT CAGCTCACACAGGAGGAGAGATGGAAGAAGGCTCCTT 2558

QY 2714 GAGTTGGAGGGTCTACCA CCACTACATCCAGGCAGCTGGAGGTTACATGGTCTCTTGCAT 2773

Db 2559 GAGTTGGAGGGTCTACCA CCACTACATCCAGGCAGCTGGAGGTTACATGGTCTCTTGCAT 2618

QY 2774 AATTTCTTCTTCTGTTGTTG TGTGATCGTCTTCTTAACGATCTTACGTTCTTGGTGGCTGAG 2833

Db 2619 AATTTCTTCTTCTGTTGTTG TGTGATCGTCTTCTTAACGATCTTACGTTCTTGGTGGCTGAG 2678

QY 2834 CTACTGTTGGAGCAGGGCTCG GGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATGGC 2893

Db 2679 CTACTGTTGGAGCAGGGCTCG GGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATGGC 2738

QY	2894	AGACCTGGGCAACATTGCAGACAAATCCTCAACTGTCTCCTTACCAGCTGGTGTACGGGCT	2953
Db	2739	AGACCTGGGCAACATTGCAGACAAATCCTCAACTGTCTCCTTACCAGCTGGTGTACGGGCT	2798
QY	2954	CAACGCCCTGCTCCTCATCTGTGTGGGGGTCTGCTCCTCAGGGATTTTCACCAAAAGTCAC	3013
Db	2799	CAACGCCCTGCTCCTCATCTGTGTGGGGGTCTGCTCCTCAGGGATTTTCACCAAAAGTCAC	2858
QY	3014	GAGGAAGGCATCCACGGCCCTGCACAACAAGTCTTCAACAAGTTTTCGCCCTGCCCCAT	3073
Db	2859	GAGGAAGGCATCCACGGCCCTGCACAACAAGTCTTCAACAAGTTTTCGCCCTGCCCCAT	2918
QY	3074	GAGTTTCTTTGACACCATCCCAATAGGCCGGCTTTTGAAGTCTTCGACGGGACTTGGG	3133
Db	2919	GAGTTTCTTTGACACCATCCCAATAGGCCGGCTTTTGAAGTCTTCGACGGGACTTGGG	2978
QY	3134	ACAGCTGGACCAAGCTCTTGCCCATCTTTTCAGAGCAGTTTCCCTGGTCCCTTAATGGT	3193
Db	2979	ACAGCTGGACCAAGCTCTTGCCCATCTTTTCAGAGCAGTTTCCCTGGTCCCTTAATGGT	3038
QY	3194	GATCGCCGTCCTGTGATGTGCAGTGTCTCTCCATATATCCTGTTAATGGGAGCCAT	3253
Db	3039	GATCGCCGTCCTGTGATGTGCAGTGTCTCTCCATATATCCTGTTAATGGGAGCCAT	3098
QY	3254	AATCATGGTTATTGCTTTCATTTATTATAT-----	3283
Db	3099	AATCATGGTTATTGCTTTCATTTATTATATGTGAGTAGGTTCTTTTTCGTTGTGACTTGG	3158
QY	3284	-----	3283
Db	3159	GGAGCAAGGGCTGGGACCAACCCAGACTAGATGGTCCCAGAGGTGGACGGTCCAGGTCCC	3218
QY	3284	-----GATGTTCAAGAAAGGCCATCGGTGTGTCAAGAGACTGGA	3322
Db	3219	TTACCTCCACTGTCCATGCAGGATGTTCAAGAAAGGCCATCGGTGTGTCAAGAGACTGGA	3278
QY	3323	GAACATAGCCGGTCTCCTTTATTCTCCACATCCTCAATTCTCTGCAAGGCCCTGAGCTC	3382
Db	3279	GAACATAGCCGGTCTCCTTTATTCTCCACATCCTCAATTCTCTGCAAGGCCCTGAGCTC	3338
QY	3383	CATCCATGTCTATGGAAAACTGAAGACTTCATCAGCCAG-----	3422
Db	3339	CATCCATGTCTATGGAAAACTGAAGACTTCATCAGCCAGTGAAGTCTCTGCTGCCATT	3398
QY	3423	-----	3422
Db	3399	TTGAGAAATGATGGAACCAACCAGGGGTGGGTGGGGAGCCAGGGAAGAAATGGGACGTCTTG	3458
QY	3423	-----	3422
Db	3459	AGAGTGGATCATCTTCAAAAAGCATTCAGAGAGCCACATTGGTTCGATTGAGACGTATTCT	3518
QY	3423	-----TTTAAGAGGCTGACTGATGCGCAG	3446
Db	3519	CTGAGCCTTCCAGAAACCTGCTGGAACCAATTCTTTCAATTAAGAGGCTGACTGATGCGCAG	3578
QY	3447	AATAACTACCTGCTGTTGTTTCTATCTTCACACGATGGATGGCATTGAGGCTGGAGATC	3506
Db	3579	AATAACTACCTGCTGTTGTTTCTATCTTCACACGATGGATGGCATTGAGGCTGGAGATC	3638
QY	3507	ATGACCAACCTTGTGACCTTGGCTGTTGCCCTGTTTCGTGGCTTTTGGCATTTCTCCACC	3566
Db	3639	ATGACCAACCTTGTGACCTTGGCTGTTGCCCTGTTTCGTGGCTTTTGGCATTTCTCCACC	3698
QY	3567	CCCTACTCCTTTAAAGTCATGGCTGTCAACATCGTGTGCAGCTGGCGTCCAGCTTCCAG	3626
Db	3699	CCCTACTCCTTTAAAGTCATGGCTGTCAACATCGTGTGCAGCTGGCGTCCAGCTTCCAG	3758
QY	3627	GCCACTGCCCGGATTTGGCTTGGAGACAGAGGCACAGTTTCACGGCTGTAGAGAGGATACTG	3686
Db	3759	GCCACTGCCCGGATTTGGCTTGGAGACAGAGGCACAGTTTCACGGCTGTAGAGAGGATACTG	3818
QY	3687	CAGTACATGAAGATGTGTGCTCGGAAGCTCCTTTACACATGGAAGGCACAAAGTTGTCCC	3746

Db	3819	CAGTACATGAAGATGTGTCTCGGAAGCTCCTTTACACATGGAAGGCACAAGTTGTCCC	3878
QY	3747	CAGGGGTGCCACAGCATGGGGAATCATATTTCAGGATTATCACATGAATAACAGAGAC	3806
Db	3879	CAGGGGTGCCACAGCATGGGGAATCATATTTCAGGATTATCACATGAATAACAGAGAC	3938
QY	3807	AACACACCCACCGTCTTCACGGCATCAACCTGACCATCCGGCGCCACGAAGTGGTGGC	3866
Db	3939	AACACACCCACCGTCTTCACGGCATCAACCTGACCATCCGGCGCCACGAAGTGGTGGC	3998
QY	3867	ATCGTGGGAAGGACGGGCTCTTGGGAAGTCTTCTTGGGCATGGCTCTCTCCGCCCTGGTG	3926
Db	3999	ATCGTGGGAAGGACGGGCTCTTGGGAAGTCTTCTTGGGCATGGCTCTCTCCGCCCTGGTG	4058
QY	3927	GAGCCCATGGCAGGCCGGATTCTCATTTAGCGGCTGGACATTTGCAGCATCGGCCCTGGAG	3986
Db	4059	GAGCCCATGGCAGGCCGGATTCTCATTTAGCGGCTGGACATTTGCAGCATCGGCCCTGGAG	4118
QY	3987	GACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAGATCCAGTCTGCTCTCAGGAACCATC	4046
Db	4119	GACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAGATCCAGTCTGCTCTCAGGAACCATC	4178
QY	4047	AGATTCAACCTAGATCCCTTTTGACCGTCAACTGACCAGCAGATCTGGGATGCCCTTGGAG	4106
Db	4179	AGATTCAACCTAGATCCCTTTTGACCGTCAACTGACCAGCAGATCTGGGATGCCCTTGGAG	4238
QY	4107	AGGACATTCCTGACCAAGGCCATCTCAAAGTTCCCCAAAAAGCTGCATACAGATGTGGTG	4166
Db	4239	AGGACATTCCTGACCAAGGCCATCTCAAAGTTCCCCAAAAAGCTGCATACAGATGTGGTG	4298
QY	4167	GAANAACGGTGGAAACTTCTCTGTGGGGGAGAGCAGCTGCTCTGCATTTGCCAGGGCTGTG	4226
Db	4299	GAANAACGGTGGAAACTTCTCTGTGGGGGAGAGCAGCTGCTCTGCATTTGCCAGGGCTGTG	4358
QY	4227	CTTCGCAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCTCCATTGACATGGAGACA	4286
Db	4359	CTTCGCAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCTCCATTGACATGGAGACA	4418
QY	4287	GACACCCGTATCCAGCGCACAAATCCGTGAAGCCTTCCAGGGCTGCACCGTCTGTCATT	4346
Db	4419	GACACCCGTATCCAGCGCACAAATCCGTGAAGCCTTCCAGGGCTGCACCGTCTGTCATT	4478
QY	4347	GCCACCCGTGTCAACCATGTGCTGAACCTGTGACCACATCCTGGTTATGGCAATGGGAAG	4406
Db	4479	GCCACCCGTGTCAACCATGTGCTGAACCTGTGACCACATCCTGGTTATGGCAATGGGAAG	4538
QY	4407	GTGGTAGAATTTGATCGGCCGGAGGTACTCGCGAAGAAAGCCTGGGTTCATTGTCGACGC	4466
Db	4539	GTGGTAGAATTTGATCGGCCGGAGGTACTCGCGAAGAAAGCCTGGGTTCATTGTCGACGC	4598
QY	4467	CTCATGGCCACAGCCACTTCTTTCACCTGAGATTAGGAGATGTGGAGACTTCATGGAGGCTG	4526
Db	4599	CTCATGGCCACAGCCACTTCTTTCACCTGAGATAAGGAGATGTGGAGACTTCATGGAGGCTG	4658
QY	4527	GCAGCTGAGCTCAGAGGTTTACACAGGTGCAGCTTCGAGGCCACAGTCTGCGACCTTCT	4586
Db	4659	GCAGCTGAGCTCAGAGGTTTACACAGGTGCAGCTTCGAGGCCACAGTCTGCGACCTTCT	4718
QY	4587	TGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAAATGTAGGGGGCTGGGGATTGC	4616
Db	4719	TGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAAATGTAGGGGGCTGGGGATTGC	4778
QY	4647	TGGATGGAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCAAGACC	4706
Db	4779	TGGATGGAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCAAGACC	4838
QY	4707	ATCTAAGACATGGGATTCAGTGATCATGTGGTCTCTCTTTTAACATTACATGCTGAATAAT	4766
Db	4839	ATCTAAGACATGGGATTCAGTGATCATGTGGTCTCTCTTTTAACATTACATGCTGAATAAT	4898
QY	4767	TTTATAAATAAGGTAAAAAGCTTATAGTTTCTTGATCTGTGTAGAGTGTTCGAAATGCTG	4826

Db	4899	TTTATAATAAGGTAAAGCTTATAGTTTCTGATCTGTGTAGAAGTGTGGCAAATGCTG	4958
QY	4827	TACTGACTTTGTAAAAATATAAACTAAGGAAAACTC	4862
Db	4959	TACTGACTTTGTAAAAATATAAACTAAGGAAAACTC	4994
RESULT 5			
AAD36023			
ID	AAD36023	standard; cDNA; 4074 BP.	
XX			
AC	AAD36023;		
XX			
DT	09-AUG-2002	(first entry)	
XX			
DE	Human adenosine triphosphate-binding cassette protein cDNA #1.		
XX			
KW	Human; adenosine triphosphate-binding cassette; chemoprotectant; ATP;		
KW	ABC protein; gene therapy; breast proliferative fibrocystic disease;		
KW	breast adenocarcinoma; drug resistance; breast disorder; transgenic;		
KW	vaccine; gene; ABCP; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..3996	
FT		/*tag= a	
FT		/product= "Human ABC protein"	
XX			
PN	WO200224742-A2.		
XX			
PD	28-MAR-2002.		
XX			
PF	20-SEP-2001; 2001WO-US29455.		
XX			
PR	22-SEP-2000; 2000US-0668628.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
XX			
PI	Krasnow RE, Baughn MR;		
XX			
DR	WPI; 2002-394129/42.		
DR	P-PSDB; AAE22722.		
XX			
PT	Mammalian adenosine triphosphate-binding cassette protein, and encoding		
PT	cDNA, useful for diagnosis and treatment of breast disorders, e.g.		
PT	breast proliferative fibrocystic disease and breast adenocarcinoma -		
XX			
PS	Claim 2; Fig 1; 80pp; English.		
XX			
CC	The invention relates to adenosine triphosphate (ATP)-binding cassette		
CC	(ABC) proteins (ABCP) and nucleic acid molecules encoding such proteins.		
CC	Sequences of the invention are useful for the diagnosis and treatment of		
CC	breast disorders, particularly breast proliferative fibrocystic disease,		
CC	breast adenocarcinoma and drug resistance. Polynucleotides of the		
CC	invention are used to produce transgenic cell lines or organisms which		
CC	are model systems for human breast disorders. They are also used in gene		
CC	therapy. ABCP sequences are useful in screening assays of phagemid or		
CC	B-lymphocyte immunoglobulin library to identify antibodies having desired		
CC	specificity. They are also useful as chemoprotectants and as vaccines.		
CC	The present sequence is a cDNA encoding human ABC protein.		
XX			
SQ	Sequence 4074 BP; 961 A; 1050 C; 1080 G; 983 T; 0 other;		
	Query Match	81.5%; Score 3964.2; DB 24; Length 4074;	
	Best Local Similarity	99.8%; Pred. No. 0;	
	Matches 3969; Conservative	0; Mismatches 8; Indels 0; Gaps 0;	
QY	601	AGCCCCCTGGACAATGCTGGCCTGTTCTCCTACCTCACCGTGTCAATGGCTCACCCCGCTCA	660
Db	98	ATCCCCCTGGACAATGCTGGCCTGTTCTCCTACCTCACCGTGTCAATGGCTCACCCCGCTCA	157
QY	661	TGATCCAAAGCTTACGGAGTCGCTTAGATGAGACACCATCCCTCCACTGTCAGTCCATG	720

Db	158	TGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTGTCAGTCCATG	217
QY	721	ATGCCTCAGACAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGTCTCAAGGC	780
Db	218	ATGCCTCAGACAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGTCTCAAGGC	277
QY	781	GAGGGATTGAAAAAGCTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGAACAAAGTTGA	840
Db	278	GAGGGATTGAAAAAGCTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGAACAAAGTTGA	337
QY	841	TTTTCGATGCACCTTCTGGGCATCTGCTTCTGCAATTGCCAGTGACTCGGGCCAATATTGA	900
Db	338	TTTTCGATGCACCTTCTGGGCATCTGCTTCTGCAATTGCCAGTGACTCGGGCCAATATTGA	397
QY	901	TTATACCAAAGATCCTGGAATATTCAGAAAGAGCAGTTGGGGAATGTTGTCCATGGAGTGG	960
Db	398	TTATACCAAAGATCCTGGAATATTCAGAAAGAGCAGTTGGGGAATGTTGTCCATGGAGTGG	457
QY	961	GACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTCTCCTCCAGTT	1020
Db	458	GACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTCTCCTCCAGTT	517
QY	1021	GGATCATCAACCAACGCACAGCCCATCAGGTTCCGAGCAGCTGTTTCTCCTTTGCCTTTG	1080
Db	518	GGATCATCAACCAACGCACAGCCCATCAGGTTCCGAGCAGCTGTTTCTCCTTTGCCTTTG	577
QY	1081	AGAAGCTCATCCAAATTTAAAGTCTGTAATACACATCACTCAGGAGAGGCCATCAGCTCT	1140
Db	578	AGAAGCTCATCCAAATTTAAAGTCTGTAATACACATCACTCAGGAGAGGTAATCAGCTTCT	637
QY	1141	TCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTAGTACTGATCA	1200
Db	638	TCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTAGTACTGATCA	697
QY	1201	CCTGCGCATCGTGGTCACTGTCAGCATTTCTTCTCTACTTCATTATTGGATACACTGCAT	1260
Db	698	CCTGCGCATCGTGGTCACTGTCAGCATTTCTTCTCTACTTCATTATTGGATACACTGCAT	757
QY	1261	TTATTGCCATCTTATGCTATCTCCTGGTTTTTCCCCTGCGGGTATTTCATGACACAAGTGG	1320
Db	758	TTATTGCCATCTTATGCTATCTCCTGGTTTTTCCCCTGCGGGTATTTCATGACACAAGTGG	817
QY	1321	CTGTGAAGGCTCAGCATCACACATCTGAGGTGAGGACCAAGCGCATCCGTGTGACCAGTG	1380
Db	818	CTGTGAAGGCTCAGCATCACACATCTGAGGTGAGGACCAAGCGCATCCGTGTGACCAGTG	877
QY	1381	AAGTTCTCACTTGCAATTAAGCTGATTAAAAATGTACATGGGAGAAAAACCAATTTGCAAAA	1440
Db	878	AAGTTCTCACTTGCAATTAAGCTGATTAAAAATGTACATGGGAGAAAAACCAATTTGCAAAA	937
QY	1441	TCATTGAAGACCTAAGAAGGAAGGAAAGAACTATTGGAGAAGTGGGGCTTGTCCAGA	1500
Db	938	TCATTGAAGACCTAAGAAGGAAGGAAAGAACTATTGGAGAAGTGGGGCTTGTCCAGA	997
QY	1501	GCCTGACAAAGTATAACCTTGTTTCATCATCCCCACAGTGGCCACAGCGGTCTGGGTTCTCA	1560
Db	998	GCCTGACAAAGTATAACCTTGTTTCATCATCCCCACAGTGGCCACAGCGGTCTGGGTTCTCA	1057
QY	1561	TCCACACATCCTTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATGCTGGCCTCCT	1620
Db	1058	TCCACACATCCTTTAAAGCTGAAACTCGCAGCGTCAATGGCCTTCAGCATGCTGGCCTCCT	1117
QY	1621	TGAATCTCCTTCGGCTGTCAGTGTTCCTTTGTGCCCTATTGCAGTCAAAGGTCACGCAATT	1680
Db	1118	TGAATCTCCTTCGGCTGTCAGTGTTCCTTTGTGCCCTATTGCAGTCAAAGGTCACGCAATT	1177
QY	1681	CCAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTTCTCCAGGAGAGCCCTGTTTCTATG	1740
Db	1178	CTAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTTCTCCAGGAGAGCCCTGTTTCTATG	1237
QY	1741	TCCAGACATTTACAGACCCCCAGCAAAAGCTCTGGTCTTTGAGGAGGCCACCTTGTCATGGC	1800

Db	1238	TCCAGACATTACAAGACCCCCAGCAAAAGCTCTGGTCTTTGAGGAGGCCACCTTGTCAATGGC	1297
QY	1801	AACAGACCTGTCCCGGGATCGTCAATGGGGGACTGGAGCTGGAGAGGAACGGGCATGCTT	1860
Db	1298	AACAGACCTGTCCCGGGATCGTCAATGGGGGACTGGAGCTGGAGAGGAACGGGCATGCTT	1357
QY	1861	CTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGGCCAGAGGAAGAAGGGAACAGCCTGG	1920
Db	1358	CTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGGCCAGAGGAAGAAGGGAACAGCCTGG	1417
QY	1921	GCCCAGAGTTGCACAAAGATCAACCTGGTGTGTCCAAAGGGGATGATGTTAGGGGTCTGCG	1980
Db	1418	GCCCAGAGTTGCACAAAGATCAACCTGGTGTGTCCAAAGGGGATGATGTTAGGGGTCTGCG	1477
QY	1981	GCAACACGGGGAGTGGTAAGAGCAGCCTGTGTTCAGCCATCCTGGAGGAGATGCACTTGC	2040
Db	1478	GCAACACGGGGAGTGGTAAGAGCAGCCTGTGTTCAGCCATCCTGGAGGAGATGCACTTGC	1537
QY	2041	TCGAGGGCTCGGTGGGGGTGCAGGGAAGCCTGGCCCTATGTCCCCCAGCAGGCCTGGATCG	2100
Db	1538	TCGAGGGCTCGGTGGGGGTGCAGGGAAGCCTGGCCCTATGTCCCCCAGCAGGCCTGGATCG	1597
QY	2101	TCAGCGGGAACATCAGGGAGAAACATCCTCATGCGGAGGCGCATATGACAAGGCCCGATACC	2160
Db	1598	TCAGCGGGAACATCAGGGAGAAACATCCTCATGCGGAGGCGCATATGACAAGGCCCGATACC	1657
QY	2161	TCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACTTCTGCCCCTTGGAGACA	2220
Db	1658	TCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACTTCTGCCCCTTGGAGACA	1717
QY	2221	TGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGCGGGGCAGAAACAGAGGATCAGCC	2280
Db	1718	TGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGCGGGGCAGAAACAGAGGATCAGCC	1777
QY	2281	TGGCCCGCGCGCTCTATTCCGACCGTCAAGATCTACCTGCTGGACGACCCCTGTCTGCTG	2340
Db	1778	TGGCCCGCGCGCTCTATTCCGACCGTCAAGATCTACCTGCTGGACGACCCCTGTCTGCTG	1837
QY	2341	TGGACGCCCCACGTGGGGAAGCACATTTTGTAGGAGTGCATTAAGAAACACTCAGGGGGA	2400
Db	1838	TGGACGCCCCACGTGGGGAAGCACATTTTGTAGGAGTGCATTAAGAAACACTCAGGGGGA	1897
QY	2401	AGACGGTCGTCTGGTGACCCACCGCTGCAGTACTTAGAAATTTTGTGCCCAGATCATTT	2460
Db	1898	AGACGGTCGTCTGGTGACCCACCGCTGCAGTACTTAGAAATTTTGTGCCCAGATCATTT	1957
QY	2461	TGTTGGAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTAATGCAGAAAAGG	2520
Db	1958	TGTTGGAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTAATGCAGAAAAGG	2017
QY	2521	GGAAATATGCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGG	2580
Db	2018	GGAAATATGCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGG	2077
QY	2581	ACACAGCAAAGATAGCAGAGAAGCCAAAGGTAGAAAAGTCAGGCTCTGGCCACCTCCCTGG	2640
Db	2078	ACACAGCAAAGATAGCAGAGAAGCCAAAGGTAGAAAAGTCAGGCTCTGGCCACCTCCCTGG	2137
QY	2641	AAGAGTCTCTCAACGGAAAATGCTGTGCCGGAGCATCAGCTCACACAGGAGGAGATGG	2700
Db	2138	AAGAGTCTCTCAACGGAAAATGCTGTGCCGGAGCATCAGCTCACACAGGAGGAGATGG	2197
QY	2701	AAGAAGGCTCCTTGAGTTGGAGGGTCTACCACCACTACATCCAGGCAGCTGGAGGTTACA	2760
Db	2198	AAGAAGGCTCCTTGAGTTGGAGGGTCTACCACCACTACATCCAGGCAGCTGGAGGTTACA	2257
QY	2761	TGGTCTCTTGCATAAATTTCTTCTTCGTGGTGTGATCGTCTTCTTAACGATCTTCAGCT	2820
Db	2258	TGGTCTCTTGCATAAATTTCTTCTTCGTGGTGTGATCGTCTTCTTAACGATCTTCAGCT	2317
QY	2821	TCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATFAGCAGCCGAGAGAGCA	2880
Db	2318	TCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATFAGCAGCCGAGAGAGCA	2377

QY	2881	ATGGAACCATGCGAGACCTGGGCAACATTGCGAGACAATCCTCAACTGTCTTCTACCAGC	2940
Db	2378	ATGGAACCATGCGAGACCTGGGCAACATTGCGAGACAATCCTCAACTGTCTTCTACCAGC	2437
QY	2941	TGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCCTCAGGGATTT	3000
Db	2438	TGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCCTCAGGGATTT	2497
QY	3001	TCACCAAAGTCACGAGGAAGGCATCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTT	3060
Db	2498	TCACCAAAGTCACGAGGAAGGCATCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTT	2557
QY	3061	TCCGCTGCCCATGAGTTTCTTTGACACCATCCCCAANTAGGCCGGCTTTTGAACGTCTCG	3120
Db	2558	TCCGCTGCCCATGAGTTTCTTTGACACCATCCCCAANTAGGCCGGCTTTTGAACGTCTCG	2617
QY	3121	CAGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCCATCTTTTCAGAGCAGTTCTTGGTCC	3180
Db	2618	CAGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCCATCTTTTCAGAGCAGTTCTTGGTCC	2677
QY	3181	TGTCCCTTAATGGTGATCGCCGTCTCTGTGATTGTTCAGTGTGCTGTCTCCATATATCCTGT	3240
Db	2678	TGTCCCTTAATGGTGATCGCCGTCTCTGTGATTGTTCAGTGTGCTGTCTCCATATATCCTGT	2737
QY	3241	TAATGGGAGCCATAATCATGGTTATTTTGCCTTCATTTATTATATGATGTTCAGAAGGCCA	3300
Db	2738	TAATGGGAGCCATAATCATGGTTATTTTGCCTTCATTTATTATATGATGTTCAGAAGGCCA	2797
QY	3301	TCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCTCTTTATTCCTCCACATCCTCA	3360
Db	2798	TCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCTCTTTATTCCTCCACATCCTCA	2857
QY	3361	ATTCTCTGCAAGGCCCTGAGCTCCAJCCATGCTTATGGAAAACTGAAGACTTCATCAGCC	3420
Db	2858	ATTCTCTGCAAGGCCCTGAGCTCCAJCCATGCTTATGGAAAACTGAAGACTTCATCAGCC	2917
QY	3421	AGTTTAAGAGGCTGACTGATGCGCAGAAATAACTACCTGCTGTTGTTTCTATCTTCCACAC	3480
Db	2918	AGTTTAAGAGGCTGACTGATGCGCAGAAATAACTACCTGCTGTTGTTTCTATCTTCCACAC	2977
QY	3481	GATGGATGGCATTGAGGCTGGAGATCATCACCACCTTGTGACCTTGGCTGTTGCCCTGT	3540
Db	2978	GATGGATGGCATTGAGGCTGGAGATCATCACCACCTTGTGACCTTGGCTGTTGCCCTGT	3037
QY	3541	TCGTGGCTTTTGGCATTTCTCTCCACCCCTACTCTCTTTAAAGTCATGGCTGTCAACATCG	3600
Db	3038	TCGTGGCTTTTGGCATTTCTCTCCACCCCTACTCTCTTTAAAGTCATGGCTGTCAACATCG	3097
QY	3601	TGCTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCCCGGATTTGGCTTGGAGACAGAGGCAC	3650
Db	3098	TGCTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCCCGGATTTGGCTTGGAGACAGAGGCAC	3157
QY	3661	AGTTCACGGCTGTAGAGAGGATACTGCACTACATGAAGATGTGTCTCGGAAGCTCCTT	3720
Db	3158	AGTTCACGGCTGTAGAGAGGATACTGCACTACATGAAGATGTGTCTCGGAAGCTCCTT	3217
QY	3721	TACACATGGAAGGCACAAGTTGTCCCGAGGGTGGCCACAGCATGGGAAAAATCATATTTC	3780
Db	3218	TACACATGGAAGGCACAAGTTGTCCCGAGGGTGGCCACAGCATGGGAAAAATCATATTTC	3277
QY	3781	AGGATTATCACATGAAATACAGAGACAACACACCCACCTGCTTACGGCATCAACCTGA	3840
Db	3278	AGGATTATCACATGAAATACAGAGACAACACACCCACCTGCTTACGGCATCAACCTGA	3337
QY	3841	CCATCCCGGCCACGAAGTGGTGGGCATCGTGGGAAGGACGGGCTCTGCGGAAGTCTCCT	3900
Db	3338	CCATCCCGGCCACGAAGTGGTGGGCATCGTGGGAAGGACGGGCTCTGCGGAAGTCTCCT	3397
QY	3901	TGGGCATGGCTCTCTTCCGCCCTGGTGGAGCCCATGGCAGGCCGGAATCTCATTTGACGGCG	3960
Db	3398	TGGGCATGGCTCTCTTCCGCCCTGGTGGAGCCCATGGCAGGCCGGAATCTCATTTGACGGCG	3457

QY 3961 TGGACATTGACAGCATCGGCCCTGGAGGACTTGGCGGTCCAAGCTCTCAGTGATCCCTCAAG 4020
|||||
Db 3458 TGGACATTGACAGCATCGGCCCTGGAGGACTTGGCGGTCCAAGCTCTCAGTGATCCCTCAAG 3517

QY 4021 ATCCAGTGTGCTCTCAGGAACCATCATCAGATTCAACCTAGATCCCTTTGACCGTCACACTG 4080
|||||
Db 3518 ATCCAGTGTGCTCTCAGGAACCATCATCAGATTCAACCTAGATCCCTTTGACCGTCACACTG 3577

QY 4081 ACCAGCAGATCTGGGATGCCCTGGAGAGGACATTCCTGACCAAGGCCATCTCAAAAGTTCC 4140
|||||
Db 3578 ACCAGCAGATCTGGGATGCCCTGGAGAGGACATTCCTGACCAAGGCCATCTCAAAAGTTCC 3637

QY 4141 CCRAAAAGCTGCATACAGATGTGGTGGAAACGGTGGAAACTTCTCTGTGGGGGAGAGGC 4200
|||||
Db 3638 CCRAAAAGCTGCATACAGATGTGGTGGAAACGGTGGAAACTTCTCTGTGGGGGAGAGGC 3697

QY 4201 AGCTGCTCTGCATTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCCTTATCGATGAAG 4260
|||||
Db 3698 AGCTGCTCTGCATTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCCTTATCGATGAAG 3757

QY 4261 CCACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATCCGTGAAGCCT 4320
|||||
Db 3758 CCACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATCCGTGAAGCCT 3817

QY 4321 TCCAGGGCTGCACCGTGTCTGTCATTGCCCAACCGTGTACCCACTGTGCTGAACGTGACC 4380
|||||
Db 3818 TCCAGGGCTGCACCGTGTCTGTCATTGCCCAACCGTGTACCCACTGTGCTGAACGTGACC 3877

QY 4381 ACATCCTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCACAGCCACTTCTTCACTGAGATAAG 4440
|||||
Db 3878 GCATCCTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCAGGTAAGTACTGCGGA 3937

QY 4441 AGAAGCCTGGGTCAATTGTTGCGAGCCCTCATGGCCACAGCCACTTCTTCACTGAGATAAG 4500
|||||
Db 3938 AGAAGCCTGGGTCAATTGTTGCGAGCCCTCATGGCCACAGCCACTTCTTCACTGAGATAAG 3997

QY 4501 GAGATGTGGAGACTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTTCACACAGGTGCAGCT 4560
|||||
Db 3998 GAGATGTGGAGACTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTTCACACAGGTGCAGCT 4057

QY 4561 TCGAGGCCACACAGTCTG 4577
|||||
Db 4058 TCGAGGCCACACAGTCTG 4074

RESULT 6
AAF83637

ID AAF83637 standard; cDNA; 3189 BP.

XX
AC AAF83637;

XX
DT 23-JUL-2001 (first entry)

XX
DE Novel human transporter protein (NHP) encoding cDNA.

XX
KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
gene therapy; ss.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers
FT CDS 1..3189
FT /*tag= a

XX
PN WO200132706-A2.

XX
PD 10-MAY-2001.

XX
PF 31-OCT-2000; 2000WO-US29852.

XX
PR 02-NOV-1999; 99US-0163018.

XX
PA (LEXI-) LEXICON GENETICS INC.

XX

PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;

XX

DR WPI; 2001-343477/36.

DR P-PSDB; AAB62549.

XX

PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT .

XX

PS Disclosure; Page 29-30; 59pp; English.

XX

CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAF83635-647 represent NHP nucleic acid sequences.

XX

SQ Sequence 3189 BP; 760 A; 828 C; 867 G; 734 T; 0 other;

Query Match 65.5%; Score 3185.8; DB 22; Length 3189;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3187; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1308 ATGACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTACGACCGCATC 1367

Db 1 ATGACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTACGACCGCATC 60

QY 1368 CGTGTGACCAGTGAAGTTCTCACTTGCAATTAAGCTGATTAATAATGTACACATGGGAGAAA 1427

Db 61 CGTGTGACCAGTGAAGTTCTCACTTGCAATTAAGCTGATTAATAATGTACACATGGGAGAAA 120

QY 1428 CCATTTGCAAAAATCAATGAAGACCTAAGAAGGAAGAAAGAAACTATTGGAGAAATGC 1487

Db 121 CCATTTGCAAAAATCAATGAAGACCTAAGAAGGAAGAAAGAAAGCTATTGGAGAAATGC 180

QY 1488 GGGCTTGTCCAGAGCCCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCG 1547

Db 181 GGGCTTGTCCAGAGCCCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCG 240

QY 1548 GTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGC 1607

Db 241 GTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGC 300

QY 1608 ATGCTGGCCTCCTTGAATCTCCTTCGGCTGTCAAGTGTCTTGTGCCTATTGCAGTCAAA 1667

Db 301 ATGCTGGCCTCCTTGAATCTCCTTCGGCTGTCAAGTGTCTTGTGCCTATTGCAGTCAAA 360

QY 1668 GGTCTCACGAATTCGAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTCCTCCAGGAGAGC 1727

Db 361 GGTCTCACGAATTCGAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTCCTCCAGGAGAGC 420

QY 1728 CCTGTTTTCTATGTCAGACATTAACAAGACCCCAAGAAAGCTCTGGTCTTTGAGGAGGCC 1787

Db 421 CCTGTTTTCTATGTCAGACATTAACAAGACCCCAAGAAAGCTCTGGTCTTTGAGGAGGCC 480

QY 1788 ACCTTGTTCATGGCAACAGACACCTGTCCCGGGGATCGTCAATGGGGCACTGGAGCTGGAGAGG 1847

Db 481 ACCTTGTTCATGGCAACAGACACCTGTCCCGGGGATCGTCAATGGGGCACTGGAGCTGGAGAGG 540

QY 1848 AACGGGCATGCTTCTGAGGGGATGACCAAGCCCTACAGATGCCCTCGGGCCACAGGAAGAA 1907

Db 541 AACGGGCATGCTTCTGAGGGGATGACCAAGCCCTACAGATGCCCTCGGGCCACAGGAAGAA 600

QY 1908 GGGAAACAGCCTGGGCCACAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAAGGGGATGATG 1967

|||||

Db	601	GGGAACAGCCTGGCCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATG	660
QY	1968	TTAGGGGTCTGCGGCAACACAGGGGAGTGGTAAGAGCAGCCCTGTTGTACCCATCCTGGAG	2027
Db	661	TTAGGGGTCTGCGGCAACACAGGGGAGTGGTAAGAGCAGCCCTGTTGTACCCATCCTGGAG	720
QY	2028	GAGATGCACTTGCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCTGGCTATGTCCCCCAG	2087
Db	721	GAGATGCACTTGCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCTGGCTATGTCCCCCAG	780
QY	2088	CAGGCCTGGATCCCTCAGCGGGAACATCAGGAGAAACATCCTCATGGAGGCGCATATGAC	2147
Db	781	CAGGCCTGGATCCCTCAGCGGGAACATCAGGAGAAACATCCTCATGGAGGCGCATATGAC	840
QY	2148	AAGGCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGACCTTGGAACTTCTG	2207
Db	841	AAGGCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGACCTTGGAACTTCTG	900
QY	2208	CCCTTTGGAGACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGGCAGAAA	2267
Db	901	CCCTTTGGAGACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGGCAGAAA	960
QY	2268	CAGAGGATCAGCCTGGCCCCCGCCGCTCTATTCCGACCCGTCAGATCTACCTGCTGGACGAC	2327
Db	961	CAGAGGATCAGCCTGGCCCCCGCCGCTCTATTCCGACCCGTCAGATCTACCTGCTGGACGAC	1020
QY	2328	CCCTGTCTGCTGTGGACGCCACCGTGGGGAAGCACATTTTGGAGAGTGCATTAAAGAAG	2387
Db	1021	CCCTGTCTGCTGTGGACGCCACCGTGGGGAAGCACATTTTGGAGAGTGCATTAAAGAAG	1080
QY	2388	ACACTCAGGGGAAGACGGTCGTCCTGCTGACCCACCAGCTGCAGTACTTAGAATTTGT	2447
Db	1081	ACACTCAGGGGAAGACGGTCGTCCTGCTGACCCACCAGCTGCAGTACTTAGAATTTGT	1140
QY	2448	GGCCAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAATCAGTGTAGTTA	2507
Db	1141	GGCCAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAATCAGTGTAGTTA	1200
QY	2508	ATGCAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAGGAAGCCACTTCG	2567
Db	1201	ATGCAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAGGAAGCCACTTCG	1260
QY	2568	GACATGTTGCAGGACACAGCAAAGATAGCAGAGAAGCCAAAGGTAGAAAGTCAGGCTCTG	2627
Db	1261	GACATGTTGCAGGACACAGCAAAGATAGCAGAGAAGCCAAAGGTAGAAAGTCAGGCTCTG	1320
QY	2628	GCCACCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCCGGAGCATCAGCTCACACAG	2687
Db	1321	GCCACCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCCGGAGCATCAGCTCACACAG	1380
QY	2688	GAGGAGGAGATGGAAGAAGGCTCCTTGAGTTGGAGGGTCTACCCACACTACATCCAGGCA	2747
Db	1381	GAGGAGGAGATGGAAGAAGGCTCCTTGAGTTGGAGGGTCTACCCACACTACATCCAGGCA	1440
QY	2748	GCTGGAGGTTACATGGTCTCTTGCATAATTTTCTTCTCGTGGTGCTGATCGTCTTCTTA	2807
Db	1441	GCTGGAGGTTACATGGTCTCTTGCATAATTTTCTTCTCGTGGTGCTGATCGTCTTCTTA	1500
QY	2808	ACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGC	2867
Db	1501	ACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGC	1560
QY	2868	AGCCGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTTGCAGACAAATCCTCAACTG	2927
Db	1561	AGCCGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTTGCAGACAAATCCTCAACTG	1620
QY	2928	TCCTTCTACCAAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGC	2987
Db	1621	TCCTTCTACCAAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGC	1680
QY	2988	TCCTCAGGGATTTTCACCAAAGTCACGAGGAAGGCATCCACGGGCCCTGCACACAAGCTC	3047
Db	1681	TCCTCAGGGATTTTCACCAAAGTCACGAGGAAGGCATCCACGGGCCCTGCACACAAGCTC	1740

QY	3048	TTCAACAAGGTTTTCGGCTGCCCCCATGAGTTTCTTTGACACCATCCCAATAGCCCGCTT	3107
Db	1741	TTCAACAAGGTTTTCGGCTGCCCCCATGAGTTTCTTTGACACCATCCCAATAGCCCGCTT	1800
QY	3108	TTGAACTGCTTCGCAGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAG	3167
Db	1801	TTGAACTGCTTCGCAGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAG	1860
QY	3168	CAGTTCCCTGGTCTGTCCCTTAATGGTGATCGCCGTCCTGTGATTGTCAAGTGTGCTGT	3227
Db	1861	CAGTTCCCTGGTCTGTCCCTTAATGGTGATCGCCGTCCTGTGATTGTCAAGTGTGCTGT	1920
QY	3228	CCATATATCCTGTTAATGGGAGCCATAATCATGGTTATTTGCTTCAATTTATATATGATG	3287
Db	1921	CCATATATCCTGTTAATGGGAGCCATAATCATGGTTATTTGCTTCAATTTATATATGATG	1980
QY	3288	TTCAAGAAAGGCATCGGTGTGTTCAAAGAGACTGGAGAACTATAGCCGCTCTCCTTTATTC	3347
Db	1981	TTCAAGAAAGGCATCGGTGTGTTCAAAGAGACTGGAGAACTATAGCCGCTCTCCTTTATTC	2040
QY	3348	TCCCACATCCTCAATTCTCTGCAAGGCCCTGAGCTCCATCCATGTCTATGGAANAAC TGAA	3407
Db	2041	TCCCACATCCTCAATTCTCTGCAAGGCCCTGAGCTCCATCCATGTCTATGGAANAAC TGAA	2100
QY	3408	GACTTCATCAGCCAGTTTAAAGAGGCTGACTGATGCGCAGAAATAACTACCTGTGTGTTT	3467
Db	2101	GACTTCATCAGCCAGTTTAAAGAGGCTGACTGATGCGCAGAAATAACTACCTGTGTGTTT	2160
QY	3468	CTATCTTCCACACGATGGATGGCATTTGAGGCTGGAGATCATGACCAACCTGTGACCTTG	3527
Db	2161	CTATCTTCCACACGATGGATGGCATTTGAGGCTGGAGATCATGACCAACCTGTGACCTTG	2220
QY	3528	GCTGTTGCCCTGTTGCTGGCTTTTGGCATTTTCCACCCCTACTCCTTTAAAGTCATG	3587
Db	2221	GCTGTTGCCCTGTTGCTGGCTTTTGGCATTTTCCACCCCTACTCCTTTAAAGTCATG	2280
QY	3588	GCTGTCAACATCGTCTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCGCCGATTTGGCTTG	3647
Db	2281	GCTGTCAACATCGTCTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCGCCGATTTGGCTTG	2340
QY	3648	GAGACAGAGGCACAGTTCACGGCTGTAGAGAGGATACTGCAGTACATGAAGATGTGTGTC	3707
Db	2341	GAGACAGAGGCACAGTTCACGGCTGTAGAGAGGATACTGCAGTACATGAAGATGTGTGTC	2400
QY	3708	TCGGAAGCTCCTTTACACATGGAAGGCACAAAGTTGTCCCCAGGGGTGGCCACAGCATGGG	3767
Db	2401	TCGGAAGCTCCTTTACACATGGAAGGCACAAAGTTGTCCCCAGGGGTGGCCACAGCATGGG	2460
QY	3768	GAAATCATATTTCAGGATTTATCATGAAATACAGAGACAACACACCCACCGTGTTCAC	3827
Db	2461	GAAATCATATTTCAGGATTTATCATGAAATACAGAGACAACACACCCACCGTGTTCAC	2520
QY	3828	GGCATCAACCTGACCATCCGCGGCCACGAAGTGGTGGGCATCGTGGAAAGGACGGCTCT	3887
Db	2521	GGCATCAACCTGACCATCCGCGGCCACGAAGTGGTGGGCATCGTGGAAAGGACGGCTCT	2580
QY	3888	GGGAAGTCCCTCTTGGGCATGGCTCTCTTCCGCCCTGGTGGAGCCCATGGCAGGCCGATTT	3947
Db	2581	GGGAAGTCCCTCTTGGGCATGGCTCTCTTCCGCCCTGGTGGAGCCCATGGCAGGCCGATTT	2640
QY	3948	CTCATTTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAGGACTTTGGGTCCAAGCTCTCA	4007
Db	2641	CTCATTTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAGGACTTTGGGTCCAAGCTCTCA	2700
QY	4008	GTGATCCCTCAAGATCCAGTGTGCTCTCAGGAACCATCAGATTTCAACCTAGATCCCTTT	4067
Db	2701	GTGATCCCTCAAGATCCAGTGTGCTCTCAGGAACCATCAGATTTCAACCTAGATCCCTTT	2760
QY	4068	GACCGTCACACTGACCAGCAGATCTGGGATGCCCTGGAGAGGACATTCCTGACCAAGGCC	4127
Db	2761	GACCGTCACACTGACCAGCAGATCTGGGATGCCCTGGAGAGGACATTCCTGACCAAGGCC	2820

QY	3351	CACATCCTCAATTCTCTGCAAGGCTGAGCTCCATCCATGTCATATGGAATAACTGAAGAC	3410
Db	2512		
QY	3411	TTCATCAGCCAGTTTAAGAGGCTGACTGATGCGCAGATAACTACCTGCTGTTGTTCTA	3470
Db	2572		
QY	3471	TCCTCCACACGATGGATGGCATTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT	3530
Db	2632		
QY	3531	GTTGCCCTGTTTCGTGGCTTTTGGCATTTCCTCCACCCCTACTCCTTTAAAGTCATGGCT	3590
Db	2692		
QY	3591	GTCAACATCGTGTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCCGGATTGGCTTGGAG	3650
Db	2752		
QY	3651	ACAGAGGCACAGTTACAGGCTGTAGAGAGGATACTGCAGTACATGAAGATGTGTCTCG	3710
Db	2812		
QY	3711	GAAGTCCTTTACACATGGAAGGCACAAAGTTGTCCCCAGGGTGGCCACAGCATGGGAA	3770
Db	2872		
QY	3771	ATCATATTTCAGGATTATCATATGAATAACAGAGACACACACCCACCGTGTTCACGGC	3830
Db	2932		
QY	3831	ATCAACCTGACCATCCCGGGCCACGAAGTGGTGGGCATCGTGGGAAGGACGGGCTCTGGG	3890
Db	2992		
QY	3891	AAGTCCTCTTGGGCATGGCTCTCTCCGCCCTGGTGGAGCCCATGGCAGGCCGATTCCTC	3950
Db	3052		
QY	3951	ATTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAGGACTTGGGTCCAAAGCTCTCAGTG	4010
Db	3112		
QY	4011	ATCCCTCAAGATCCAGTGCTGCTCTCAGGAACCATCAGATTCAAACCTAGATCCCTTTGAC	4070
Db	3172		
QY	4071	CGTCACACTGACCAGCAGATCTGGGATGCCCTTGGAGAGGACATTCCTGACCAAGCCATC	4130
Db	3232		
QY	4131	TCAAAAGTTCCCCAAAAAGCTGCATACAGATGTGGTGGAAAAACGGTGGAAACTTCTCTGTG	4190
Db	3292		
QY	4191	GGGAGAGGCAGCTGCTCTGCATTCGAGGGCTGTGCTTCGCAACTCCAAGATCATCCTT	4250
Db	3352		
QY	4251	ATCGATGAAGCCACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATC	4310
Db	3412		
QY	4311	CGTGAAGCCTTCCAGGGCTGCACCGTGTCTCGTCATTGGCCACCGTGTCAACCACTGTGCTG	4370
Db	3472		
QY	4371	AACTGTGACCACATCCTGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCGGAG	4430
Db	3532		

QY	4431	GTA	CTACTGCGGAAGACCTGGGTCA	TGTTGCGAGCCCTCATGGCCACAGCCACTTCTTCA	4490
Db	3592	GTA	CTACTGCGGAAGACCTGGGTCA	TGTTGCGAGCCCTCATGGCCACAGCCACTTCTTCA	3651
QY	4491	CTG	AGATAA	4499	
Db	3652	CTG	AGATAA	3660	
RESULT 8					
AAF83642					
ID	AAF83642 standard; cDNA; 3075 BP.				
AC	AAF83642;				
XX	23-JUL-2001 (first entry)				
DE	Novel human transporter protein (NHP) encoding cDNA.				
XX	Novel human protein; transporter protein; NHP; therapeutic; diagnostic;				
KW	gene therapy; ss.				
XX	Homo sapiens.				
OS	Location/Qualifiers				
XX	Key				
FT	CDS				
FT	1..3075				
FT	/*tag= a *				
FT	misc_feature				
FT	259..267				
FT	/*tag= b				
FT	/note= "the amino acid residues encoded by the above 3				
FT	codons are not indicated in the corresponding				
FT	protein sequence"				
XX	WO200132706-A2.				
PN	10-MAY-2001.				
XX	31-OCT-2000; 2000WO-US29852.				
PF	02-NOV-1999; 99US-0163018.				
XX	(LEXI-) LEXICON GENETICS INC.				
PA	Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;				
XX	Sands AT;				
PI	WPI; 2001-343477/36.				
XX	P-PSDB; AAB62554.				
DR	Novel isolated human polynucleotide sequences encoding polypeptides				
XX	that share sequence similarity with mammalian multidrug resistance				
PT	proteins and cellular transporter proteins, useful as probe or primer				
PS	Disclosure; Page 41-42; 59pp; English.				
XX	The invention relates to novel human transporter proteins (NHP) and				
CC	polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences				
CC	can be used in conjunction with PCR to screen libraries, isolate clones				
CC	and prepare cloning and sequencing templates. The NHP oligonucleotides				
CC	can also be used as hybridization probes for screening libraries, for				
CC	assessing gene patterns and for preparing antisense nucleic acid				
CC	molecules. The NHP nucleotide sequences are also useful in screening				
CC	techniques for drugs which treats symptomatic or phenotypic				
CC	manifestations of perturbing the normal function of NHP in the body.				
CC	Sequences AAF83635-647 represent NHP nucleic acid sequences.				
XX	Sequence 3075 BP; 739 A; 784 C; 803 G; 749 T; 0 other;				
SQ					

Query Match 63.1%; Score 3066.6; DB 22; Length 3075;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3069; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 351 ATGACTAGGAAGAGGACATACTGGGTGCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 410
Db 1 ATGACTAGGAAGAGGACATACTGGGTGCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 60
QY 411 ATCGACATAGGCGATGACATGGTTTCAGGACTTAATTATAAAACCTATACTCTCCAAGAT 470
Db 61 ATCGACATAGGCGATGACATGGTTTCAGGACTTAATTATAAAACCTATACTCTCCAAGAT 120
QY 471 GGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGCAGCTGTGCCACCG 530
Db 121 GGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGCAGCTGTGCCACCG 180
QY 531 TGGGGGAAGTATGATGCTGCCCTTGAGAAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT 590
Db 181 TGGGGGAAGTATGATGCTGCCCTTGAGAAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT 240
QY 591 CCTGCCCCCCCAGCCCCCTGGACAAATGCTGGCCTGTTCTCTCTACCTCACCCGTGTCATGGCTC 650
Db 241 CCTGCCCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTCTCTACCTCACCCGTGTCATGGCTC 300
QY 651 ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAAACACCATCCCTCCACTG 710
Db 301 ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAAACACCATCCCTCCACTG 360
QY 711 TCAGTCCATGATGCCCTCAGACAAAATGTCCAAAGCTTCACCGCCTTTGGGAAGAA 770
Db 361 TCAGTCCATGATGCCCTCAGACAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAA 420
QY 771 GTCTCAAGGCGAGGATGAAAAAGCTTCAGTGCCTCTGCTGATGCTGAGGTTCCAGAGA 830
Db 421 GTCTCAAGGCGAGGATGAAAAAGCTTCAGTGCCTCTGCTGATGCTGAGGTTCCAGAGA 480
QY 831 ACAAGGTTGATTTTCGATGCACCTCTGGGCATCTGCTTCTGCAATGGCCAGTGTACTCGG 890
Db 481 ACAAGGTTGATTTTCGATGCACCTCTGGGCATCTGCTTCTGCAATGGCCAGTGTACTCGG 540
QY 891 CCAATATTGATTATACCAAGATCCTGGAATATTCAGAAGAGCAGTTGGGSAATGTTGTC 950
Db 541 CCAATATTGATTATACCAAGATCCTGGAATATTCAGAAGAGCAGTTGGGSAATGTTGTC 600
QY 951 CATGGAGTGGGACTCIGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC 1010
Db 601 CATGGAGTGGGACTCIGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC 660
QY 1011 TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCAAGCAGCTGTTCTCTCC 1070
Db 661 TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCAAGCAGCTGTTCTCTCC 720
QY 1071 TTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC 1130
Db 721 TTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC 780
QY 1131 ATCAGCTTCTTCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGACCCCTA 1190
Db 781 ATCAGCTTCTTCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGACCCCTA 840
QY 1191 GTACTGATCACCTGCGCATCGCTGGTCATCTGCAGCATTTCTTCCCTACTTCATTATTGGA 1250
Db 841 GTACTGATCACCTGCGCATCGCTGGTCATCTGCAGCATTTCTTCCCTACTTCATTATTGGA 900
QY 1251 TACACTGCATTTATTGCCATCTTATGCTATCTCTGGTTTTCACCTGGCGGTATTTCATG 1310
Db 901 TACACTGCATTTATTGCCATCTTATGCTATCTCTGGTTTTCACCTGGAGGTATTTCATG 960
QY 1311 ACAAGAAATGGCTGTAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT 1370
Db 961 ACAAGAAATGGCTGTAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT 1020
QY 1371 GTGACCAGTGAAGTTCTCACTTGCAATTAGCTGATTAANAATGTACACATGGGAGAAACCA 1430
Db 1021 GTGACCAGTGAAGTTCTCACTTGCAATTAGCTGATTAANAATGTACACATGGGAGAAACCA 1080
QY 1431 TTTGCAAAAATCATTTGAAGACCTTAAGAAGGAAGGAAAGAACTATTGGAGAGTGC 1490

Db 1081 TTTTGCAAAAATCATTTGAAGACCTTAAGAAAGGAAGGAAGCTATTGGAGAAAGTCGGGG 1140
QY 1491 CTTGTCCAGAGCCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGGTC 1550
Db 1141 CTTGTCCAGAGCCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGGTC 1200
QY 1551 TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG 1610
Db 1201 TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG 1260
QY 1611 CTGGCCTCCTTGAATCTCCTTCGGCTGTCACTGTCTTGTGCCCTATTGCAGTCAAAAGGT 1670
Db 1261 CTGGCCTCCTTGAATCTCCTTCGGCTGTCACTGTCTTGTGCCCTATTGCAGTCAAAAGGT 1320
QY 1671 CTCACGAATTCCAAGTCTGCAGTGTAGAGGTTCAAGAAAGTTTTCCTCCAGGAGAGCCCT 1730
Db 1321 CTCACGAATTCCAAGTCTGCAGTGTAGAGGTTCAAGAAAGTTTTCCTCCAGGAGAGCCCT 1380
QY 1731 GTTTTCTATGTCCAGACATTTACAAGACCCCCCAGCAAAAGCTCTTGGTCTTTGAGGAGGCCACC 1790
Db 1381 GTTTTCTATGTCCAGACATTTACAAGACCCCCCAGCAAAAGCTCTGGTCTTTGAGGAGGCCACC 1440
QY 1791 TTGTTCATGGCAACAGACCTGTCCCCGGGATCGTCAATGGGGCAGTGGAGCTGGAGAGGAAC 1850
Db 1441 TTGTTCATGGCAACAGACCTGTCCCCGGGATCGTCAATGGGGCAGTGGAGCTGGAGAGGAAC 1500
QY 1851 GGGCATGCTTCTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGGG 1910
Db 1501 GGGCATGCTTCTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGGG 1560
QY 1911 AACAGCCTGGGCCCAGAGTTGCACAAAGATCAACCTGGTGTGCCAAGGGGATGATGTTA 1970
Db 1561 AACAGCCTGGGCCCAGAGTTGCACAAAGATCAACCTGGTGTGCCAAGGGGATGATGTTA 1620
QY 1971 GGGGTCTGCGGCAACACGGGGAGTGGTAAGAGCAGAGCCTGTGTCAGCCATCCTGGAGGAG 2030
Db 1621 GGGGTCTGCGGCAACACGGGGAGTGGTAAGAGCAGAGCCTGTGTCAGCCATCCTGGAGGAG 1680
QY 2031 ATGCACCTTGTCTGAGGGCTCGGTGGGGTGCAGGGAAGCCTTGGCCTATGTCCCCCAGCAG 2090
Db 1681 ATGCACCTTGTCTGAGGGCTCGGTGGGGTGCAGGGAAGCCTTGGCCTATGTCCCCCAGCAG 1740
QY 2091 GCCTGGATCGTCAAGGGGAACATCAGGGGAGAACATCCTCATGGGAGCGGCATATGACAAG 2150
Db 1741 GCCTGGATCGTCAAGGGGAACATCAGGGGAGAACATCCTCATGGGAGCGGCATATGACAAG 1800
QY 2151 GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACCTTCTGCCC 2210
Db 1801 GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTTCTGCCC 1860
QY 2211 TTTGGAGACATGACAGATTGGAGAGCGGGCCCTCAACCTCTCTGGGGGCGCAAAACAG 2270
Db 1861 TTTGGAGACATGACAGATTGGAGAGCGGGCCCTCAACCTCTCTGGGGGCGCAAAACAG 1920
QY 2271 AGGATCAGCCTGGCCCGCCCGCTCTATTCCGACCGTCAGATCTACCTGCTGGAGCACCC 2330
Db 1921 AGGATCAGCCTGGCCCGCCCGCTCTATTCCGACCGTCAGATCTACCTGCTGGAGCACCC 1980
QY 2331 CTGCTGCTGTGGACGCCACCGTGGGGAAGCACATTTTGGAGGAGTGCAATTAAAGAGACA 2390
Db 1981 CTGCTGCTGTGGACGCCACCGTGGGGAAGCACATTTTGGAGGAGTGCAATTAAAGAGACA 2040
QY 2391 CTCAGGGGGAAGACGGTCTGCTCCTGGTGACCCACAGCTGCAGTACTTAGAATTTGTGGC 2450
Db 2041 CTCAGGGGGAAGACGGTCTGCTCCTGGTGACCCACAGCTGCAGTACTTAGAATTTGTGGC 2100
QY 2451 CAGATCATTTTGTGGAAAAATGGGAAAAATCTGTGAAAAATGGAACCTCACAGTCACTTAATG 2510
Db 2101 CAGATCATTTTGTGGAAAAATGGGAAAAATCTGTGAAAAATGGAACCTCACAGTCACTTAATG 2160
QY 2511 CAGAAAAAGGGGAAATPATGCCCAACTTATCCAGAAAGATGCACAAGGAAGCCACTTCGGAC 2570

Db 2161 CAGAAAAAGGGGAAATATATGCCCAACTTATCCAGAAGATGCACAAGGAGCCACTTCGGAC 2220

QY 2571 ATGTTGCAGGACACAGCAAAAGATAGCAGAGAAGCCAAAAGGTAGAAAAGTCAGGCTCTGGCC 2630

Db 2221 ATGTTGCAGGACACAGCAAAAGATAGCAGAGAAGCCAAAAGGTAGAAAAGTCAGGCTCTGGCC 2280

QY 2631 ACCTCCCTGGAAGAGTCTCTCAACGGGAAATGCTGTCCGGAGCATCAGTCACACAGGAG 2690

Db 2281 ACCTCCCTGGAAGAGTCTCTCAACGGGAAATGCTGTCCGGAGCATCAGTCACACAGGAG 2340

QY 2691 GAGGAGATGGAAGAGCTCCTTGAGTTGGAGGGTCTACCAACCACCTACATCCAGGCAGCT 2750

Db 2341 GAGGAGATGGAAGAGGCTCCTTGAGTTGGAGGGTCTACCAACCACCTACATCCAGGCAGCT 2400

QY 2751 GGAGGTTACATGGTCTCTTGCATAAATTTTCTTCTCGTGGTCTGATCGTCTTTAAGC 2810

Db 2401 GGAGGTTACATGGTCTCTTGCATAAATTTTCTTCTCGTGGTCTGATCGTCTTTAAGC 2460

QY 2811 ATCTTCAGCTTCGTGGCTGAGCTACTGTTGGAGCAGGGCTCGGGGACCAATAGCAGC 2870

Db 2461 ATCTTCAGCTTCGTGGCTGAGCTACTGTTGGAGCAGGGCTCGGGGACCAATAGCAGC 2520

QY 2871 CGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTGCAGACAATCCTCAACTGTCC 2930

Db 2521 CGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTGCAGACAATCCTCAACTGTCC 2580

QY 2931 TTCTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGTCTCC 2990

Db 2581 TTCTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGTCTCC 2640

QY 2991 TCAGGGATTTTCACCAAAAGTCACGAGGAGGAGCATCCACGGCCCTGCACAACAAGCTCTTC 3050

Db 2641 TCAGGGATTTTCACCAAAAGTCACGAGGAGGAGCATCCACGGCCCTGCACAACAAGCTCTTC 2700

QY 3051 AACAAAGGTTTCCGCTGCCCATGAGTTCTTTGACACCATCCCAATAGGCCGGCTTTTG 3110

Db 2701 AACAAAGGTTTCCGCTGCCCATGAGTTCTTTGACACCATCCCAATAGGCCGGCTTTTG 2760

QY 3111 AACTGCTTCGAGGGGACTTGGAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAGCAG 3170

Db 2761 AACTGCTTCGAGGGGACTTGGAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAGCAG 2820

QY 3171 TTCCTGGTCCCTCCTTAATGGTGATCGCGCTCCTGTTGATTGTCAAGTGTCTGTCTCCA 3230

Db 2821 TTCCTGGTCCCTCCTTAATGGTGATCGCGCTCCTGTTGATTGTCAAGTGTCTGTCTCCA 2880

QY 3231 TATATCCTGTTAATGGGAGCCATAATCATGTTATTTGCTTCATTTATATATATGATGTTT 3290

Db 2881 TATATCCTGTTAATGGGAGCCATAATCATGTTATTTGCTTCATTTATATATATGATGTTT 2940

QY 3291 AAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTCTCC 3350

Db 2941 AAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTCTCC 3000

QY 3351 CACATCCTCAATTCTCTGCAAGGCCCTGAGCTCCATCCATGCTCTATGGAAAAACTGAAGAC 3410

Db 3001 CACATCCTCAATTCTCTGCAAGGCCCTGAGCTCCATCCATGCTCTATGGAAAAACTGAAGAC 3060

QY 3411 TTCATCAGCCAGT 3423

Db 3061 TTCATCAGCCAGT 3073

RESULT 9
ABV21036
ID ABV21036 standard; cDNA; 3055 BP.
XX
AC ABV21036;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 21027.
XX

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
OS
XX WO200160860-A2.
PN
XX 23-AUG-2001.
PD
XX 20-FEB-2001; 2001WO-US05171.
PF
XX 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
PI
XX WPI; 2001-662795/76.
DR
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 3467; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 3055 BP; 731 A; 765 C; 845 G; 712 T; 2 other;

Query Match 61.0%; Score 2967.2; DB 23; Length 3055;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 3006; Conservative 0; Mismatches 3; Indels 27; Gaps 1;

QY 1824 AATGGGGCACTGGAGCTGGAGAGGAACCGGCATGCTTCTGAGGGGATGACCGCCTAGA 1883

Db 27 AATGGGGCACTGGAGCTGGAGAGGAACCGGCATGCTTCTGAGGGGATGACCGCCTAGA 86

QY 1884 GATGCCCTCGGCCAGAGGAAGAAAGGGAACAGCCTGGCCCCAGAGTTGCACAAGATCAAC 1943

Db 87 GATGCCCTCGGG-----CCAGAGTTGCACAAGATCAAC 119

QY 1944 CTGGTGGTGTCCAAAGGGGATGATGTTAGGGGTCTCGGCAACACGGGAGTGGTAAGAGC 2003

Db 120 CTGGTGGTGTCCAAAGGGGATGATGTTAGGGGTCTCGGCAACACGGGAGTGGTAAGAGC 179

QY 2004 AGCCTGTGTCAGCCATCCTCGGAGGAGATGCACATTGCTCGAGGGCTCGGTGGGGTGCAG 2063

Db 180 AGCCTGTGTCAGCCATCCTCGGAGGAGATGCACATTGCTCGAGGGCTCGGTGGGGTGCAG 239

QY 2064 GGAAGCCTGGCCTATGTCCCCCAGCAGGCCTGGATCGTCAGCGGGGAACATCAGGGAGAAC 2123

Db 240 GGAAGCCTGGCCTATGTCCCCCAGCAGGCCTGGATCGTCAGCGGGGAACATCAGGGAGAAC 299

QY 2124 ATCCTCATGGGAGCGGCATATGACAAGCCCCGATACCTCCAGGTGCTCCACTGCTGCTCC 2183

Db 300 ATCCTCATGGAGGCGCATATGACAAGGCCCGATACCTCCAGGTGCTCCACTGCTGCTCC 359

QY 2184 CTGAATCGGACCTGGAACTTCTGCCCTTTGGAGACATGACAGAGATTGGAGCGGGGC 2243

Db 360 CTGAATCGGACCTGGAACTTCTGCCCTTTGGAGACATGACAGAGATTGGAGCGGGGC 419

QY 2244 CTCAACCTCTCTGGGGGCGAGAAACAGAGGATCAGCCTGGCCCCGCCGCTCTATTCCGAC 2303

Db 420 CTCAACCTCTCTGGGGGCGAGAAACAGAGGATCAGCCTGGCCCCGCCGCTCTATTCCGAC 479

QY 2304 CGTCAGATCTACCTGCTGGACGACCCCTGTCTGTCTGTGGACGCCACGTTGGGAAGCAC 2363

Db 480 CGTCAGATCTACCTGCTGGACGACCCCTGTCTGTCTGTGGACGCCACGTTGGGAAGCAC 539

QY 2364 ATTTTGGAGAGTGCAATTAAGAAGACACTCAGGGGGAAGACGGTCTCTGTTGACCCAC 2423

Db 540 ATTTTGGAGAGTGCAATTAAGAAGACACTCAGGGGGAAGACGGTCTCTGTTGACCCAC 599

QY 2424 CAGCTGCAGTACTTAGAATTTTGTGGCCAGATCATTTTGTGGAAATGGGAAATCTGT 2483

Db 600 CAGCTGCAGTACTTAGAATTTTGTGGCCAGATCATTTTGTGGAAATGGGAAATCTGT 659

QY 2484 GAAATGGAACCTCACAGTGAGTTAATGCAGAAAAAGGGGAAATATGCCAACTTATCCAG 2543

Db 660 GAAATGGAACCTCACAGTGAGTTAATGCAGAAAAAGGGGAAATATGCCAACTTATCCAG 719

QY 2544 AAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGGACACAGCAAAAGATAGCAGAGAAG 2603

Db 720 AAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGGACACAGCAAAAGATAGCAGAGAAG 779

QY 2604 CCAAAGGTAGAAAGTCAGGGCTCTGGCCACCTCCCTCGGAAGAGTCTCTCAACGGGAAATGCT 2663

Db 780 CCAAAGGTAGAAAGTCAGGGCTCTGGCCACCTCCCTCGGAAGAGTCTCTCAACGGGAAATGCT 839

QY 2664 GTGCCGGAGCATCAGCTCACACAGGAGGAGGAGATGGAAGAAGCTCCTTGAGTTGGAGG 2723

Db 840 GTGCCGGAGCATCAGCTCACACAGGAGGAGGAGATGGAAGAAGCTCCTTGAGTTGGAGG 899

QY 2724 GTCTACCAACACTACATCCAGGCACTGGAGGTTACATGGTCTCTTGCCATATTTCTTC 2783

Db 900 GTCTACCAACACTACATCCAGGCACTGGAGGTTACATGGTCTCTTGCCATATTTCTTC 959

QY 2784 TTCTGTGTGCTGATCGTCTTCTTAACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTG 2843

Db 960 TTTGTGTGCTGATCGTCTTCTTAACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTG 1019

QY 2844 GAGCAGGCTCGGGGACCAATAGCAGCCGAGAGACAAATGGAACCATGGCAGACCTGGGC 2903

Db 1020 GAGCAGGCTCGGGGACCAATAGCAGCCGAGAGACAAATGGAACCATGGCAGACCTGGGC 1079

QY 2904 AACATTGCAGACAAATCCTCAACTGTCTTCTACAGCTGGTGTACGGGCTCAACGCCCTG 2963

Db 1080 AACATTGCAGACAAATCCTCAACTGTCTTCTACAGCTGGTGTACGGGCTCAACGCCCTG 1139

QY 2964 CTCCTCATCTGTGTGGGGGTCTGCTCCTCAGGGATTTTCACCAAGTCAACGAGGAAGCA 3023

Db 1140 CTCCTCATCTGTGTGGGGGTCTGCTCCTCAGGGATTTTCACCAAGTCAACGAGGAAGCA 1199

QY 3024 TCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTTTCGGCTGCCCATGAGTTTCTTT 3083

Db 1200 TCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTTTCGGCTGCCCATGAGTTTCTTT 1259

QY 3084 GACACCTCCCAATAGGCCGGCTTTTGAACGTCTTCGCAGGGACTTGGAACAGCTGGAC 3143

Db 1260 GACACCTCCCAATAGGCCGGCTTTTGAACGTCTTCGCAGGGACTTGGAACAGCTGGAC 1319

QY 3144 CAGCTCTTGCCCATCTTTTTCAGAGCAGTTCCTGGTCTCTTAATGGTGATCGCCGTC 3203

Db 1320 CAGCTCTTGCCCATCTTTTTCAGAGCAGTTCCTGGTCTCTTAATGGTGATCGCCGTC 1379

QY 3204 CTGTTGATGTGCTGTCTCCATATATCCTGTTAATGGGAGCCATAATCATGGTT 3263

Db 1380 CTGTTGATTTGTCAGTGTGCTGTCTCCCATATATCCTGTTAATGGGAGCCATAATCATGGTT 1439

QY 3264 ATTTGCTTCATTTATATATGATGTTCAAGAAAGGCCATCGGTGTGTTCAAGAGACTGGAG 3323

Db 1440 ATTTGCTTCATTTATATATGATGTTCAAGAAAGGCCATCGGTGTGTTCAAGAGACTGGAG 1499

QY 3324 AACTATAGCCGCTCTCCTTTATTTCTCCACATCCTCAATTCTCTGCAAGGCTGAGCTCC 3383

Db 1500 AACTATAGCCGCTCTCCTTTATTTCTCCACATCCTCAATTCTCTGCAAGGCTGAGCTCC 1559

QY 3384 ATCCATGTCTATGGAAAACTGAAGACTTTCATCAGCCAGTTTAAAGAGGCTGACTGATCG 3443

Db 1560 ATCCATGTCTATGGAAAACTGAAGACTTTCATCAGCCAGTTTAAAGAGGCTGACTGATCG 1619

QY 3444 CAGAATAACTACCTGCTGTTGTTCTATCTTCCACACGATGGATGGCATTTGAGGCTGGAG 3503

Db 1620 CAGAATAACTACCTGCTGTTGTTCTATCTTCCACACGATGGATGGCATTTGAGGCTGGAG 1679

QY 3504 ATCATGACCAACCTTGTGACCTTGGCTGTTGCCCTGTTCCGTGGCTTTTGGCATTTCTCTCC 3563

Db 1680 ATCATGACCAACCTTGTGACCTTGGCTGTTGCCCTGTTCCGTGGCTTTTGGCATTTCTCTCC 1739

QY 3564 ACCCCTACTCCTTTAAAGTCATGGCTGTCAACATCGTGTGCAGTGGCGTCCAGCTTC 3623

Db 1740 ACCCCTACTCCTTTAAAGTCATGGCTGTCAACATCGTGTGCAGTGGCGTCCAGCTTC 1799

QY 3624 CAGGCCACTGCCCCGATTGGCTTGGAGACAGAGGCACAGTTCACGGCTGTAGAGAGGATA 3683

Db 1800 CAGGCCACTGCCCCGATTGGCTTGGAGACAGAGGCACAGTTCACGGCTGTAGAGAGGATA 1859

QY 3684 CTGCAGTACATGAAGATGTGTCTCGGAAGCTCCTTTACACATGGAAGGCACAAAGTTGT 3743

Db 1860 CTGCAGTACATGAAGATGTGTCTCGGAAGCTCCTTTACACATGGAAGGCACAAAGTTGT 1919

QY 3744 CCCCAGGGTGGCCACAGCATGGGGAATCATATTTTCAGGATTTATCACATGAAATACAGA 3803

Db 1920 CCCCAGGGTGGCCACAGCATGGGGAATCATATTTTCAGGATTTATCACATGAAATACAGA 1979

QY 3804 GACAACACACCCACCTGCTTTCACGGCATCAACCTGACCATCCGCGGCCACGAAGTGGTG 3863

Db 1980 GACAACACACCCACCTGCTTTCACGGCATCAACCTGACCATCCGCGGCCACGAAGTGGTG 2039

QY 3864 GGATCCTGGGAAGGACGGGCTCTGGGAAGTCTCTTGGGCATGGCTCTCTTCCGCCCTG 3923

Db 2040 GGATCCTGGGAAGGACGGGCTCTGGGAAGTCTCTTGGGCATGGCTCTCTTCCGCCCTG 2099

QY 3924 GTGAGCCCATGGCAGGCCGATTCTCATTTGACGGCTGGGACATTTGCAGCATCGGCCCTG 3983

Db 2100 GTGAGCCCATGGCAGGCCGATTCTCATTTGACGGCTGGGACATTTGCAGCATCGGCCCTG 2159

QY 3984 GAGGACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAGATCCAGTGCTGCTCTCAGGAACC 4043

Db 2160 GAGGACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAGATCCAGTGCTGCTCTCAGGAACC 2219

QY 4044 ATCAGATTCAACTAGATCCCTTTGACCGTCAACACTGACCCAGCAGATCTGGGATGCCCTG 4103

Db 2220 ATCAGATTCAACTAGATCCCTTTGACCGTCAACACTGACCCAGCAGATCTGGGATGCCCTG 2279

QY 4104 GAGAGGACATTCCTGACCAAGGCCATCTCAAAAGTTCCCCAAAAAGCTGCATACAGATGTG 4163

Db 2280 GAGAGGACATTCCTGACCAAGGCCATCTCAAAAGTTCCCCAAAAAGCTGCATACAGATGTG 2339

QY 4164 GTGAAAAACGGTGGAAACTTCTCTGTGGGGAGAGGAGCTGCTGCTGCAATTGCCAGGCT 4223

Db 2340 GTGAAAAACGGTGGAAACTTCTCTGTGGGGAGAGGAGCTGCTGCTGCAATTGCCAGGCT 2399

QY 4224 GTGCTTCGCAACTCCAAGATCATCCTTATTCGATGAAGCCACACCTCCATTTGACATGGAG 4283

Db 2400 GTGCTTCGCAACTCCAAGATCATCCTTATTCGATGAAGCCACAGCCTCCATTTGACATGGAG 2459

QY 4284 ACAGACACCTTGATCCAGCGGCACAATCCGTGAAGCCTTCCAGGGCTGCACCGTGCCTC 4343

Db 2460 ACAGACACCTTGATCCAGCGGCACAATCCGTGAAGCCTTCCAGGGCTGCACCGTGCCTC 2519

QY 4344 ATTGCCACCCTGTCAACCACTGTGCTGAACCTGTGACCAACATCCTGGTTATGGCAATGGG 4403
|||||
Db 2520 ATTGCCACCCTGTCAACCACTGTGCTGAACCTGTGACCGCATCCTGGTTATGGCAATGGG 2579

QY 4404 AAGGTGGTAGAATTGATCGGCCGGAGGTACTGCGGGAAGAAGCCCTGGGTCAATGTTCCGA 4463
|||||
Db 2580 AAGGTGGTAGAATTGATCGGCCGGAGGTACTGCGGAAGAAGCCCTGGGTCAATGTTCCGA 2639

QY 4464 GCCCTCATGGCCACAGCCCACTTCTTCACTGAGATAAGGAGATGTGGAGACTTCATGGAGG 4523
|||||
Db 2640 GCCCTCATGGCCACAGCCCACTTCTTCACTGAGATAAGGAGATGTGGAGACTTCATGGAGG 2699

QY 4524 CTGGCAGCTGAGCTCAGAGGTTTCACACAGGTGCAGCTTCGAGGCCCAACAGTCTCGACCT 4583
|||||
Db 2700 CTGGCAGCTGAGCTCAGAGGTTTCACACAGGTGCAGCTTCGAGGCCCAACAGTCTCGACCT 2759

QY 4584 TCTTGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAAATGTAGSGGGGTGGGAT 4643
|||||
Db 2760 TCTTGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAAATGTAGSGGGGTGGGAT 2819

QY 4644 TGCTGGATGGAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCA 4703
|||||
Db 2820 TGCTGGATGGAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCA 2879

QY 4704 ACCATCTAAGACATGGGATTCAGTGATCATGTGGTTCTCCTTTTAACTTACATGCTGAAT 4763
|||||
Db 2880 ACCATCTAAGACATGGGATTCAGTGATCATGTGGTTCTCCTTTTAACTTACATGCTGAAT 2939

QY 4764 AATTTTATAATAAGTAAAGCTTATAGTTTCTGATCTGTGTTAGAAAGTGTGCAAATG 4823
|||||
Db 2940 AATTTTATAATAAGTAAAGCTTATAGTTTCTGATCTGTGTTAGAAAGTGTGCAAATG 2999

QY 4824 CTGTACTGACTTTGTAAAAATATAAACTAAGGAAAA 4859
|||||
Db 3000 CTGTACTGACTTTGTAAAAATATAAACTAAGGAAAA 3035

RESULT 10

ABV26879
ID ABV26879 standard; cDNA; 3055 BP.
XX
AC ABV26879;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 26870.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim 1; Page 5435-5436; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 3055 BP; 731 A; 765 C; 845 G; 712 T; 2 other;

Query Match 61.0%; Score 2967.2; DB 23; Length 3055;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 3006; Conservative 0; Mismatches 3; Indels 27; Gaps 1;

QY 1824 AATGGGGCACTGGAGCTGGAGAGGAACGGGCATGCTTCTGAGGGGATGACCAAGCCTAGA 1883
|||||
Db 27 AATGGGGCACTGGAGCTGGAGAGGAACGGGCATGCTTCTGAGGGGATGACCAAGCCTAGA 86

QY 1884 GATGCCCTCGGCCAGAGGAAGGAAGCAAGCCTGGGCCAGAGTGCACAAGATCAAC 1943
|||||
Db 87 GATGCCCTCGGG-----CCAGAGTTGCACAAGATCAAC 119

QY 1944 CTGGTGGTGTCAAAGGGGATGATGTTAGGGGTCTCGGGCAACACCGGGAGTGGTAAGAGC 2003
|||||
Db 120 CTGGTGGTGTCAAAGGGGATGATGTTAGGGGTCTCGGGCAACACCGGGAGTGGTAAGAGC 179

QY 2004 AGCCTGTTGTCAGCCATCCTCGGAGGAGATGCACTTGTCTCGAGGGGTCTCGTGGGGTGCAG 2063
|||||
Db 180 AGCCTGTTGTCAGCCATCCTCGGAGGAGATGCACTTGTCTCGAGGGGTCTCGTGGGGTGCAG 239

QY 2064 GGAAGCCCTGGCCTATGTCCCCCAGCAGGCGCTGGATCGTCAGCGGGGAACATCAGGGAGAAC 2123
|||||
Db 240 GGAAGCCCTGGCCTATGTCCCCCAGCAGGCGCTGGATCGTCAGCGGGGAACATCAGGGAGAAC 299

QY 2124 ATCCTCATGGAGGCGCATATGACAAAGGCCGATACCTCCAGGTGCTCCACTGCTGCTCC 2183
|||||
Db 300 ATCCTCATGGAGGCGCATATGACAAAGGCCGATACCTCCAGGTGCTCCACTGCTGCTCC 359

QY 2184 CTGAATCGGACCTGGAACTTCTGCCCTTTGGAGACATGACAGAGATGGAGAGCGGGGC 2243
|||||
Db 360 CTGAATCGGACCTGGAACTTCTGCCCTTTGGAGACATGACAGAGATGGAGAGCGGGGC 419

QY 2244 CTCAACCTCTCTGGGGGGCAGAAACAGAGGATCAGCCTGGCCCGCGCGTCTATTCCGAC 2303
|||||
Db 420 CTCAACCTCTCTGGGGGGCAGAAACAGAGGATCAGCCTGGCCCGCGCGTCTATTCCGAC 479

QY 2304 CGTCAGATCTACCTGCTGGACGACCCCTGTCTGTGTGGACGCCCACTGGGGGAAGCAC 2363
|||||
Db 480 CGTCAGATCTACCTGCTGGACGACCCCTGTCTGTGTGGACGCCCACTGGGGGAAGCAC 539

QY 2364 ATTTTGGAGAGTGCATTAAAGAAGACACTCAGGGGGGAAGACGGTCTCTCTGGTACCCAC 2423
|||||
Db 540 ATTTTGGAGAGTGCATTAAAGAAGACACTCAGGGGGGAAGACGGTCTCTCTGGTACCCAC 599

QY 2424 CAGCTGCAGTACTTAGAATTTGTGGCCAGATCATTTTGTGGAAATGGGAAATCTGT 2483
|||||
Db 600 CAGCTGCAGTACTTAGAATTTGTGGCCAGATCATTTTGTGGAAATGGGAAATCTGT 659

QY 2484 GAAAATGGAACCTCACAGTGAGTTAATGCAGAAAAAGGGGAAATATGCCCAACTTATCCAG 2543
|||||

Db	660	GAAATGGAACTCACAGTGAGTTAATGCAGAAAAAGGGAAATATATGCCCAACTTATCCAG	719
Qy	2544	AAGATGCACAAGGAAGCCACITTCGGACATGTTGCAGGACACAGCAAGATAGCAGAGAAG	2603
Db	720	AAGATGCACAAGGAAGCCACITTCGGACATGTTGCAGGACACAGCAAGATAGCAGAGAAG	779
Qy	2604	CCAAAGGTAGAAAGTCAGGCTCTGGCCACCTCCCTGGAAAGAGTCTCTCAACGGAAATGCT	2663
Db	780	CCAAAGGTAGAAAGTCAGGCTCTGGCCACCTCCCTGGAAAGAGTCTCTCAACGGAAATGCT	839
Qy	2664	GTCCCGGAGCATCAGCTCACACAGGAGGAGAGATGGAAGAAGCTCCTTGAGTTGGAGG	2723
Db	840	GTCCCGGAGCATCAGCTCACACAGGAGGAGAGATGGAAGAAGCTCCTTGAGTTGGAGG	899
Qy	2724	GTCTACCACCACTACATCCAGGCAGCTGGAGGTTACATGGTCTCTTGCCATGCTCTTC	2783
Db	900	GTCTACCACCACTACATCCAGGCAGCTGGAGGTTACATGGTCTCTTGCCATGCTCTTC	959
Qy	2784	TTCGTGTGCTGATCGTCTCTTAACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTG	2843
Db	960	TTCGTGTGCTGATCGTCTCTCTTAACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTG	1019
Qy	2844	GAGCAGGSGCTCGGGGACCAATAGCAGCCGAGAGACCAATGGAACCATGGCAGACCTGGGC	2903
Db	1020	GAGCAGGSGCTCGGGGACCAATAGCAGCCGAGAGACCAATGGAACCATGGCAGACCTGGGC	1079
Qy	2904	AACATTGCAGACAAATCCTCAACTGTCCTTCTACCAGCTGGTGTACGGGCTCAACGCCCTG	2963
Db	1080	AACATTGCAGACAAATCCTCAACTGTCCTTCTACCAGCTGGTGTACGGGCTCAACGCCCTG	1139
Qy	2964	CTCCTCATCTGTGTGGGGTCTGCTCCTCAGGGATTTTACCAAAAGTCAACGAGGAAGCA	3023
Db	1140	CTCCTCATCTGTGTGGGGTCTGCTCCTCAGGGATTTTACCAAAAGTCAACGAGGAAGCA	1199
Qy	3024	TCCACGGSCCTGCACAACAAGCTCTTCAACAAGGTTTCCGCTGCCCCATGAGTTTCTTT	3083
Db	1200	TCCACGGSCCTGCACAACAAGCTCTTCAACAAGGTTTCCGCTGCCCCATGAGTTTCTTT	1259
Qy	3084	GACACCATCCCAATAGGCGGGCTTTTGAACCTGCTTCGCAGGSGACTTGGAAACAGCTGGAC	3143
Db	1260	GACACCATCCCAATAGGCGGGCTTTTGAACCTGCTTCGCAGGSGACTTGGAAACAGCTGGAC	1319
Qy	3144	CAGCTCTTGCCCATCTTTTCAGAGCAGTTCCTGGTCCCTGCTCTTAATGGTGATCGCCGTC	3203
Db	1320	CAGCTCTTGCCCATCTTTTCAGAGCAGTTCCTGGTCCCTGCTCTTAATGGTGATCGCCGTC	1379
Qy	3204	CTGTTGATGTCAGTGTGCTGTCTCCATATATCCTGTTAATGGGAGCCATAATCATGTTT	3263
Db	1380	CTGTTGATGTCAGTGTGCTGTCTCCATATATCCTGTTAATGGGAGCCATAATCATGTTT	1439
Qy	3264	ATTTGCTTCATTTATATATGATGTTCAAGAAGGCCATCGGTGTGTTCAAGAGACTGGAG	3323
Db	1440	ATTTGCTTCATTTATATATGATGTTCAAGAAGGCCATCGGTGTGTTCAAGAGACTGGAG	1499
Qy	3324	AACTATAGCCGGTCTCCTTTATTCCTCCACATCCTCAATCTCTGCAAGGCCCTGAGCTCC	3383
Db	1500	AACTATAGCCGGTCTCCTTTATTCCTCCACATCCTCAATCTCTGCAAGGCCCTGAGCTCC	1559
Qy	3384	ATCCATGCTATGGAAAAACTGAAGACTTCATCAGCCAGTTTAAAGAGGCTGACTGATGCG	3443
Db	1560	ATCCATGCTATGGAAAAACTGAAGACTTCATCAGCCAGTTTAAAGAGGCTGACTGATGCG	1619
Qy	3444	CAGAAATACTACCTGCTGTTGTTTCTATCTCCACACGATGATGGCATTGAGGCTGGAG	3503
Db	1620	CAGAAATACTACCTGCTGTTGTTTCTATCTCCACACGATGATGGCATTGAGGCTGGAG	1679
Qy	3504	ATCATGACCAACCTTGTGACCTTGGCTGTTGCCCTGTTCCGCTTTTGGCATTTCCCTCC	3563
Db	1680	ATCATGACCAACCTTGTGACCTTGGCTGTTGCCCTGTTCCGCTTTTGGCATTTCCCTCC	1739
Qy	3564	ACCCCTACTCCTTTAAAGTCATGGCTGTCAACATCGTGTGAGCTGGCGTCCAGCTTC	3623
Db	1740	ACCCCTACTCCTTTAAAGTCATGGCTGTCAACATCGTGTGAGCTGGCGTCCAGCTTC	1799

Qy	3624	CAGGCCACTGCCCGGATTTGGCTTTGGACACAGAGGCACAGTTCACGGCTGTAGAGAGGATA	3683
Db	1800	CAGGCCACTGCCCGGATTTGGCTTTGGACACAGAGGCACAGTTCACGGCTGTAGAGAGGATA	1859
Qy	3684	CTGCAGTACATGAAGATGTGTGTCTCGGAAGCTCCTTTACACATGGAAGGCACAAAGTTGT	3743
Db	1860	CTGCAGTACATGAAGATGTGTGTCTCGGAAGCTCCTTTACACATGGAAGGCACAAAGTTGT	1919
Qy	3744	CCCCAGGGTGGCCACAGCATGGGGAATCATATTTTCAGGATTTATCACATGAAATACAGA	3803
Db	1920	CCCCAGGGTGGCCACAGCATGGGGAATCATATTTTCAGGATTTATCACATGAAATACAGA	1979
Qy	3804	GACAACACACCCACCCGTGTTTCAGGCATCAACCTGACCATCCCGGCCACGAAGTGGTG	3863
Db	1980	GACAACACACCCACCCGTGTTTCAGGCATCAACCTGACCATCCCGGCCACGAAGTGGTG	2039
Qy	3864	GGATCGTGGGAAGACGGGCTCTGGGAAGTCCCTTGGGCATGGCTCTCTCCGCCCTG	3923
Db	2040	GGATCGTGGGAAGACGGGCTCTGGGAAGTCCCTTGGGCATGGCTCTCTCCGCCCTG	2099
Qy	3924	GTGAGCCCATGGCAGCCGGATTCTCATTTGACGGCGTGGACATTTGCAGCATCGGCCCTG	3983
Db	2100	GTGAGCCCATGGCAGCCGGATTCTCATTTGACGGCGTGGACATTTGCAGCATCGGCCCTG	2159
Qy	3984	GAGGACTTTCGGTCCAAAGCTCTCAGTGATCCCTCAAGATCCAGTCTGCTCTCAGGAACC	4043
Db	2160	GAGGACTTTCGGTCCAAAGCTCTCAGTGATCCCTCAAGATCCAGTCTGCTCTCAGGAACC	2219
Qy	4044	ATCAGATTCAAACCTAGATCCCTTTGACCCGTACACTGACCAGCAGATCTGGGATGCCCTTG	4103
Db	2220	ATCAGATTCAAACCTAGATCCCTTTGACCCGTACACTGACCAGCAGATCTGGGATGCCCTTG	2279
Qy	4104	GAGAGACATTCCTGACCAAGGCCATCTCAAAAGTTCCCAAAAAGCTGCATACAGATGTG	4163
Db	2280	GAGAGACATTCCTGACCAAGGCCATCTCAAAAGTTCCCAAAAAGCTGCATACAGATGTG	2339
Qy	4164	GTGAAAAACGGTGGAAACTTCTCTGTGGGGGAGAGGCAGCTGCTGTCATTTGCCAGGGCT	4223
Db	2340	GTGAAAAACGGTGGAAACTTCTCTGTGGGGGAGAGGCAGCTGCTGTCATTTGCCAGGGCT	2399
Qy	4224	GTGCTTCGCAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCTCCAATGACATGGAG	4283
Db	2400	GTGCTTCGCAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCTCCAATGACATGGAG	2459
Qy	4284	ACAGACACCCCTGATCCAGCGCACAAATCCGTGAAGCCTTCCAGSGCTGCACCCTCGTC	4343
Db	2460	ACAGACACCCCTGATCCAGCGCACAAATCCGTGAAGCCTTCCAGSGCTGCACCCTCGTC	2519
Qy	4344	ATTGCCACCCGTGTCAACCCTGTGCTGAACCTGTGACCACATCCTGGTTATGGGCAATGGG	4403
Db	2520	ATTGCCACCCGTGTCAACCCTGTGCTGAACCTGTGACCACATCCTGGTTATGGGCAATGGG	2579
Qy	4404	AAGTGGTAGAAATTTGATCGGCCGGAGGTACTGCCGAAGAGCTGGGTTCATTTTCGCA	4463
Db	2580	AAGTGGTAGAAATTTGATCGGCCGGAGGTACTGCCGAAGAGCTGGGTTCATTTTCGCA	2639
Qy	4464	GCCCTCATGGCCACAGCCACTTCTTCACTGAGATAAGGAGATGTGGAGACTTTCATGGAGG	4523
Db	2640	GCCCTCATGGCCACAGCCACTTCTTCACTGAGATAAGGAGATGTGGAGACTTTCATGGAGG	2699
Qy	4524	CTGSCAGCTGAGCTCAGAGGTTTACACAGGTGCAGCTTCGAGGCCACACAGTCTGCGACCT	4583
Db	2700	CTGSCAGCTGAGCTCAGAGGTTTACACAGGTGCAGCTTCGAGGCCACACAGTCTGCGACCT	2759
Qy	4584	TCITGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAAATGTAGGGGGGGTGGGGAT	4643
Db	2760	TCITGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAAATGTAGGGGGGGTGGGGAT	2819
Qy	4644	TGCTGGATGGAACCCCTTGAATAGGCTACTTGTATGGCTCTCAAGACCTTAGAACCCCCAGA	4703
Db	2820	TGCTGGATGGAACCCCTTGAATAGGCTACTTGTATGGCTCTCAAGACCTTAGAACCCCCAGA	2879

QY 4704 ACCATCTAAGACATGGGATTTCAGTGATCATGTGGTTCTCCTTTTAACCTTACATGCTGAAT 4763
|||||
Db 2880 ACCATCTAAGACATGGGATTTCAGTGATCATGTGGTTCTCCTTTTAACCTTACATGCTGAAT 2939
|||||
QY 4764 AATTTATAATAAGGTAAAAGCTTATAGTTTTCGATCTGTGTTAGAAAGTGTGCAAAATG 4823
|||||
Db 2940 AATTTATAATAAGGTAAAAGCTTATAGTTTTCGATCTGTGTTAGAAAGTGTGCAAAATG 2999
|||||
QY 4824 CTGTACTGACTTTGTAAAATATATAAACTAAGGAAAA 4859
|||||
Db 3000 CTGTACTGACTTTGTAAAATATATAAACTAAGGAAAA 3035
|||||

RESULT 11
AAF83641
ID AAF83641 standard; cDNA; 2937 BP.
XX
AC AAF83641;
XX
DT 23-JUL-2001 (first entry)
XX
DE Novel human transporter protein (NHP) encoding cDNA.
XX
KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2937
FT /*tag= a
FT misc_feature 259..267
FT /*tag= b
FT /note= "the amino acid residues encoded by the above 3
FT codons are not indicated in the corresponding
FT protein sequence"
XX
PN WO200132706-A2.
XX 10-MAY-2001.
XX
PF 31-OCT-2000; 2000WO-US29852.
XX
PR 02-NOV-1999; 99US-0163018.
XX (LEXI-) LEXICON GENETICS INC.
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI; 2001-343477/36.
DR P-PSDB; AAB62553.
XX
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX
PS Disclosure; Page 38-39; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treat symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAF83635-647 represent NHP nucleic acid sequences.
XX
SQ Sequence 2937 BP; 702 A; 747 C; 775 G; 713 T; 0 other;

Query Match 60.2%; Score 2927.6; DB 22; Length 2937;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2930; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 351 ATGACTAGGAAGAGACATACTGGGTGCCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 410
|||||
Db 1 ATGACTAGGAAGAGACATACTGGGTGCCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 60
|||||
QY 411 ATCGACATAGCGGATGACATGGTTTCAGGACATATTTATAAAACCTATACTCTCCAAGAT 470
|||||
Db 61 ATCGACATAGCGGATGACATGGTTTCAGGACATATTTATAAAACCTATACTCTCCAAGAT 120
|||||
QY 471 GGCCCTGGAGTCAGCAAGAGAGAAAAATCCTGAGGCTCCAGGAGGGCAGCTGCCACCG 530
|||||
Db 121 GGCCCTGGAGTCAGCAAGAGAGAAAAATCCTGAGGCTCCAGGAGGGCAGCTGCCACCG 180
|||||
QY 531 TGGGGGAAGTATGATGCTGCCTTGAGAAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT 590
|||||
Db 181 TGGGGGAAGTATGATGCTGCCTTGAGAAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT 240
|||||
QY 591 CCTGCCCCCCAGCCCTTGACAATGCTGGCCTGTTCCTACCTCACCCGTCTCATGGCTC 650
|||||
Db 241 CCTGCCCCCCAGCCCTTGACAATGCTGGCCTGTTCCTACCTCACCCGTCTCATGGCTC 300
|||||
QY 651 ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAAACACCATCCCTCCACTG 710
|||||
Db 301 ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAAACACCATCCCTCCACTG 360
|||||
QY 711 TCAGTCCATGATGCGCTCAGACAAAAATGTCAAAGGCTTCACCGCCTTTGGGAAGAAGAA 770
|||||
Db 361 TCAGTCCATGATGCGCTCAGACAAAAATGTCAAAGGCTTCACCGCCTTTGGGAAGAAGAA 420
|||||
QY 771 GTCTCAAGGCGAGGATTTGAAAAAGCTTCAGTCTCTGGTATGCTGAGGTTCCAGAGA 830
|||||
Db 421 GTCTCAAGGCGAGGATTTGAAAAAGCTTCAGTCTCTGGTATGCTGAGGTTCCAGAGA 480
|||||
QY 831 ACAAGGTTGATTTTCGATGCACCTTCTGGGCACTCTCTGCATTGCCAGTGTACTCGGG 890
|||||
Db 481 ACAAGGTTGATTTTCGATGCACCTTCTGGGCACTCTCTGCATTGCCAGTGTACTCGGG 540
|||||
QY 891 CCAATATTGATTATACCAAAGATCCTGGAAATATTAGTCTCCGAATGTGTGAGTTTC 950
|||||
Db 541 CCAATATTGATTATACCAAAGATCCTGGAAATATTAGTCTCCGAATGTGTGAGTTTC 600
|||||
QY 951 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAGTTTC 1010
|||||
Db 601 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAGTTTC 660
|||||
QY 1011 TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCC 1070
|||||
Db 661 TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCC 720
|||||
QY 1071 TTTGCCCTTTGAGAAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC 1130
|||||
Db 721 TTTGCCCTTTGAGAAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC 780
|||||
QY 1131 ATCAGCTTCTTCACCGGTGATGTAAACTACCTGTGTTGAAGGGGTGTGCTATGGACCCCTA 1190
|||||
Db 781 ATCAGCTTCTTCACCGGTGATGTAAACTACCTGTGTTGAAGGGGTGTGCTATGGACCCCTA 840
|||||
QY 1191 GTACTGATCACCTGCGCATCGCTGGTTCATCTGCAGCATTTCTTCTCTACTTCATTATTGGA 1250
|||||
Db 841 GTACTGATCACCTGCGCATCGCTGGTTCATCTGCAGCATTTCTTCTCTACTTCATTATTGGA 900
|||||
QY 1251 TACACTGCATTATTTGCCCATCTTATGCTATCTCCTGGTTTTCCTCCTGCGGTATTTCATG 1310
|||||
Db 901 TACACTGCATTATTTGCCCATCTTATGCTATCTCCTGGTTTTCCTCCTGCGGTATTTCATG 960
|||||
QY 1311 ACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT 1370
|||||
Db 961 ACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT 1020
|||||
QY 1371 GTGACCAGTGAAGTTCTCACTTGCATTTAAGCTGATTAAATGTACACATGGGAGAAACCA 1430
|||||

FT. misc_feature 259..267
FT /*tag= b
FT /note= "the amino acid residues encoded by the above 3
FT codons are not indicated in the corresponding
FT protein sequence"
XX
PN
XX
PD
XX
XX WO200132706-A2.
PD 10-MAY-2001.
XX
PF 31-OCT-2000; 2000WO-US29852.
XX
XX 02-NOV-1999; 99US-0163018.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI; 2001-343477/36.
DR P-PSDB; AAB62556.
XX
XX Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX
XX
PS Disclosure; Page 48-49; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAF83635-647 represent NHP nucleic acid sequences.
XX
SQ Sequence 2448 BP; 582 A; 619 C; 639 G; 608 T; 0 other;

Query Match 45.5%; Score 2213.2; DB 22; Length 2448;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2215; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 351 ATGACTAGGAAGAGACATACTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 410
DB 1 ATGACTAGGAAGAGACATACTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 60

QY 411 ATCGACATAGGCGATGACATGTTTTCAGGACTTATTATAAAACCTATACTCTCCAAGAT 470
DB 61 ATCGACATAGGCGATGACATGTTTTCAGGACTTATTATAAAACCTATACTCTCCAAGAT 120

QY 471 GGCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGCAGCTGTCCCACCG 530
DB 121 GGCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGCAGCTGTCCCACCG 180

QY 531 TGGGGGAAGTATGATGCTGCCTTGAGAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT 590
DB 181 TGGGGGAAGTATGATGCTGCCTTGAGAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT 240

QY 591 CCTGCCCCCCAGCCCTGGACAAATGCTGGCCTGTTCTCCTACCTCACCGTGTATGGCTC 650
DB 241 CCTGCCCCCCAGCCCTGGACAAATGCTGGCCTGTTCTCCTACCTCACCGTGTATGGCTC 300

QY 651 ACCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTG 710
DB 301 ACCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTG 360

QY 711 TCAGTCCATGATGCCTCAGACAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA 770
DB 361 TCAGTCCATGATGCCTCAGACAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA 420

QY 771 GTCTCAAGGCGAGGGATTGAAAAAGCTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGA 830
DB 421 GTCTCAAGGCGAGGGATTGAAAAAGCTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGA 480

QY 831 ACAAGGTTGATTTTCGATGCACATTCTGGGGATCTGCTTCTGCATTGCCAGTGTACTCGG 890
DB 481 ACAAGGTTGATTTTCGATGCACATTCTGGGGATCTGCTTCTGCATTGCCAGTGTACTCGG 540

QY 891 CCAATATTGATTATACCAAAAGATCCTGGAATATTAGAAAGAGCAGTTGGGAATGTTGTC 950
DB 541 CCAATATTGATTATACCAAAAGATCCTGGAATATTAGAAAGAGCAGTTGGGAATGTTGTC 600

QY 951 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC 1010
DB 601 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC 660

QY 1011 TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCC 1070
DB 661 TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCC 720

QY 1071 TTTGCCCTTTGAGAAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC 1130
DB 721 TTTGCCCTTTGAGAAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC 780

QY 1131 ATCAGCTTCTCACCCGGTGATGTAACCTACCTGTTTGAAGGGGTGTGCTATGGACCCCTA 1190
DB 781 ATCAGCTTCTCACCCGGTGATGTAACCTACCTGTTTGAAGGGGTGTGCTATGGACCCCTA 840

QY 1191 GTACTGATCACCTGCGCATCGCTGGTGCATCTGCAGCATTTCTTCTCCTACTTCATATTGGA 1250
DB 841 GTACTGATCACCTGCGCATCGCTGGTGCATCTGCAGCATTTCTTCTCCTACTTCATATTGGA 900

QY 1251 TACACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCACTGCGGTATTCATG 1310
DB 901 TACACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCACTGCGGTATTCATG 960

QY 1311 ACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT 1370
DB 961 ACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT 1020

QY 1371 GTGACCAGTGAAGTTCTCAGTTGCATTAAGCTGATTAATAATGTACACATGGGAGAAACCA 1430
DB 1021 GTGACCAGTGAAGTTCTCAGTTGCATTAAGCTGATTAATAATGTACACATGGGAGAAACCA 1080

QY 1431 TTTGCAAAAATCATTTGAAGACCTAAGAAGCAAGGAAAGAACTATTGGAGAAGTGC GG 1490
DB 1081 TTTGCAAAAATCATTTGAAGACCTAAGAAGCAAGGAAAGAACTATTGGAGAAGTGC GG 1140

QY 1491 CTTGTCCAGAGCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGTC 1550
DB 1141 CTTGTCCAGAGCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGTC 1200

QY 1551 TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG 1610
DB 1201 TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG 1260

QY 1611 CTGGCCTCCTTGAATCTCCTTCGGCTGTCAAGTGTCTTTGTGCCCTATTGCAGTCAAAGGT 1670
DB 1261 CTGGCCTCCTTGAATCTCCTTCGGCTGTCAAGTGTCTTTGTGCCCTATTGCAGTCAAAGGT 1320

QY 1671 CTCACGAATCCCAAGTCTGCAGTGTAGGTTCAAGAAGTTTTTCTCCAGGAGAGCCCT 1730
DB 1321 CTCACGAATCCCAAGTCTGCAGTGTAGGTTCAAGAAGTTTTTCTCCAGGAGAGCCCT 1380

QY 1731 GTTTTCTATGTCCAGACATTACAAGACCCCGGATCGTCAATGGGCACTTGAGGAGGCCACC 1790
DB 1381 GTTTTCTATGTCCAGACATTACAAGACCCCGGATCGTCAATGGGCACTTGAGGAGGCCACC 1440

QY 1791 TTGTCTATGGCAACAGACCTGTCCCGGATCGTCAATGGGCACTTGAGGAGGCCAAC 1850
DB 1441 TTGTCTATGGCAACAGACCTGTCCCGGATCGTCAATGGGCACTTGAGGAGGCCAAC 1500

QY 1851 GGGCATGCTTCTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGGG 1910
|||||
Db 1501 GGGCATGCTTCTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGGG 1560

QY 1911 AACAGCCTGGGCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATGTTA 1970
|||||
Db 1561 AACAGCCTGGGCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATGTTA 1620

QY 1971 GGGGCTGCGGCAACACGCGGAGTGGTAAGAGCAGCCCTGTTGTGAGCCCATCCTGGAGGAG 2030
|||||
Db 1621 GGGGCTGCGGCAACACGCGGAGTGGTAAGAGCAGCCCTGTTGTGAGCCCATCCTGGAGGAG 1680

QY 2031 ATGCACTTGCTCGAGGGCTCGGTGGGGGTGCAGGGAAGCCTGGCCCTATGTCCCCCAGCAG 2090
|||||
Db 1681 ATGCACTTGCTCGAGGGCTCGGTGGGGGTGCAGGGAAGCCTGGCCCTATGTCCCCCAGCAG 1740

QY 2091 GCCTGGATCGTCAGCGGGAACATCAGGGGAGAACATCCTCATGGGAGGCGCATATGACAAAG 2150
|||||
Db 1741 GCCTGGATCGTCAGCGGGAACATCAGGGGAGAACATCCTCATGGGAGGCGCATATGACAAAG 1800

QY 2151 GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACTTCTGCC 2210
|||||
Db 1801 GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACTTCTGCC 1860

QY 2211 TTTGGAGACATGACAGAGATTGGAGAGCGGGCCCTCAACCTCTCTGGGGGCAGAAACAG 2270
|||||
Db 1861 TTTGGAGACATGACAGAGATTGGAGAGCGGGCCCTCAACCTCTCTGGGGGCAGAAACAG 1920

QY 2271 AGGATCAGCCTGGCCCGCGCTCTATTCCGACCGTCAGATCTACCTGCTGGACGACCCC 2330
|||||
Db 1921 AGGATCAGCCTGGCCCGCGCTCTATTCCGACCGTCAGATCTACCTGCTGGACGACCCC 1980

QY 2331 CTGCTGCTGTGGACGCCCCACGTGGGGAAAGCACATTTTGTGAGGAGTGCATTAAAGAAGACA 2390
|||||
Db 1981 CTGCTGCTGTGGACGCCCCACGTGGGGAAAGCACATTTTGTGAGGAGTGCATTAAAGAAGACA 2040

QY 2391 CTCAGGGGAAGACGGTCGTCCTGGTGACCCACCGTCGAGTACTTAGAATTTTGTGGC 2450
|||||
Db 2041 CTCAGGGGAAGACGGTCGTCCTGGTGACCCACCGTCGAGTACTTAGAATTTTGTGGC 2100

QY 2451 CAGATCATTTTGTGGAAATGGGAAAATCTGTGAAAATGGAATCAGAGTGAGTTAATG 2510
|||||
Db 2101 CAGATCATTTTGTGGAAATGGGAAAATCTGTGAAAATGGAATCAGAGTGAGTTAATG 2160

QY 2511 CAGAAAAGGGGAAATATGCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGG 2568
|||||
Db 2161 CAGAAAAGGGGAAATATGCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGG 2218

RESULT 13
AAF83645
ID AAF83645 standard; cDNA: 2586 BP.
XX
AC AAF83645;
XX
DT 23-JUL-2001 (first entry)
XX
DE Novel human transporter protein (NHP) encoding cDNA.
XX
KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2586
FT /*tag= a
FT misc_feature 259..267
FT /*tag= b
FT /note= "the amino acid residues encoded by the above 3
FT codons are not indicated in the corresponding
FT protein sequence"
XX

PN WO200132706-A2.
XX
PD 10-MAY-2001.
XX
PF 31-OCT-2000; 2000WO-US29852.
XX
PR 02-NOV-1999; 99US-0163018.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI; 2001-343477/36.
DR P-PSDB; AAB62557.
XX
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT .
XX
PS Disclosure; Page 51-52; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAF83635-647 represent NHP nucleic acid sequences.
XX
SQ Sequence 2586 BP; 619 A; 656 C; 667 G; 644 T; 0 other;

Query Match 45.5%; Score 2213.2; DB 22; Length 2586;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2215; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 351 ATGACTAGGAAGAGGACATACTAGGTGCCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 410
|||||
Db 1 ATGACTAGGAAGAGGACATACTAGGTGCCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC 60

QY 411 ATCGACATAGCGGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT 470
|||||
Db 61 ATCGACATAGCGGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT 120

QY 471 GGGCCCTGGAGTCAGCAAGAGAGAAAATCCTGAGGCTCCAGGGAGGCGAGCTGTCCACCCG 530
|||||
Db 121 GGGCCCTGGAGTCAGCAAGAGAGAAAATCCTGAGGCTCCAGGGAGGCGAGCTGTCCACCCG 180

QY 531 TGGGGGAAGTATGATGCTGCTGCTTGAGAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT 590
|||||
Db 181 TGGGGGAAGTATGATGCTGCTGCTTGAGAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT 240

QY 591 CCTGCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTCCTACCTCACCCTGTCATGGCTC 650
|||||
Db 241 CCTGCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTCCTACCTCACCCTGTCATGGCTC 300

QY 651 ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTG 710
|||||
Db 301 ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTG 360

QY 711 TCAGTCCCATGATGCCCTCAGACAAAATGTCCAAAGGCTTCACCCGCTTTGGGAAGAAGAA 770
|||||
Db 361 TCAGTCCCATGATGCCCTCAGACAAAATGTCCAAAGGCTTCACCCGCTTTGGGAAGAAGAA 420

QY 771 GTCTCAAGCGGAGGGATTGAAAAAGCTTCAGTGCCTTCTGGTGATGCTGAGGTTCCAGAGA 830
|||||
Db 421 GTCTCAAGCGGAGGGATTGAAAAAGCTTCAGTGCCTTCTGGTGATGCTGAGGTTCCAGAGA 480

QY 831 ACAAGGTTGATTTTCGATGCACTTCTTGGGCACTCTGCTGCATTTGCCAGTGTACTCGG 890

Db 1561 AACAGCCTGGGCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATGTTA 1620

QY 1971 GGGGTCTGCGGCAACACACGGGGAGTGGTAAGAGCAGCCCTGTTGTCAGCCATCCTGGAGGAG 2030

Db 1621 GGGGTCTGCGGCAACACACGGGGAGTGGTAAGAGCAGCCCTGTTGTCAGCCATCCTGGAGGAG 1680

QY 2031 ATGCAC TTGCTCGAGGGCTCGGTGGGGTGTCAGGGAAGCCCTGGCCTATGTCCCCCAGCAG 2090

Db 1681 ATGCAC TTGCTCGAGGGCTCGGTGGGGTGTCAGGGAAGCCCTGGCCTATGTCCCCCAGCAG 1740

QY 2091 GCCTGGATCGTCAGCGGGAAACATCAGGAGAGAACATCCCTCATGGAGGCGCATATGACAAG 2150

Db 1741 GCCTGGATCGTCAGCGGGAAACATCAGGAGAGAACATCCCTCATGGAGGCGCATATGACAAG 1800

QY 2151 GCCCGATACCTCCAGGTGCTCCACTGCTCCCTGAATCGGGACCTGGAAC TTCTGCC 2210

Db 1801 GCCCGATACCTCCAGGTGCTCCACTGCTCCCTGAATCGGGACCTGGAAC TTCTGCC 1860

QY 2211 TTTGGAGACATGACAGAGATTGGAGCGGGGCCCTCAACCTCTCTGGGGGCGAGAAACAG 2270

Db 1861 TTTGGAGACATGACAGAGATTGGAGCGGGGCCCTCAACCTCTCTGGGGGCGAGAAACAG 1920

QY 2271 AGGATCAGCCTGGCCCCGCCGCTCTATTCCGACCGTCAGATCTACCTGCTGGACGACCCC 2330

Db 1921 AGGATCAGCCTGGCCCCGCCGCTCTATTCCGACCGTCAGATCTACCTGCTGGACGACCCC 1980

QY 2331 CTGCTGCTGTGGACGCCCCACGCTGGGGAAGCACATTTTGGAGAGTGCA TTAAAGAAGACA 2390

Db 1981 CTGCTGCTGTGGACGCCCCACGCTGGGGAAGCACATTTTGGAGAGTGCA TTAAAGAAGACA 2040

QY 2391 CTCAGSGGGAAGACGGTCTGCTCCTGGTGACCCACCAGCTGCAGTACTTAGAATTTTGTGGC 2450

Db 2041 CTCAGSGGGAAGACGGTCTGCTCCTGGTGACCCACCAGCTGCAGTACTTAGAATTTTGTGGC 2100

QY 2451 CAGATCATTTTGTGGAAAATCTGTGAAAATGGAATGGAAC TACAGTGAGTTAATG 2510

Db 2101 CAGATCATTTTGTGGAAAATCTGTGAAAATGGAATGGAAC TACAGTGAGTTAATG 2160

QY 2511 CAGAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGG 2568

Db 2161 CAGAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGG 2218

RESULT 14
AAF83640
ID AAF83640 standard; cDNA; 2700 BP.

XX AAF83640;

XX 23-JUL-2001 (first entry)

XX Novel human transporter protein (NHP) encoding cDNA.

DE Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; ss.

XX Homo sapiens.

OS Key Location/Qualifiers
FH CDS 1..2700
FT /*tag= a

XX WO200132706-A2.

XX 10-MAY-2001.

XX 31-OCT-2000; 2000WO-US29852.

XX 02-NOV-1999; 99US-0163018.

XX (LEXI-) LEXICON GENETICS INC.

PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;

PI Sands AT;
XX
DR WPI: 2001-343477/36.
DR P-PSDB; AAB62552.
XX
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT -
XX
XX
PS Disclosure; Page 36; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs. fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides
CC can also be used as hybridization probes for screening libraries, for
CC assessing gene patterns and for preparing antisense nucleic acid
CC molecules. The NHP nucleotide sequences are also useful in screening
CC techniques for drugs which treats symptomatic or phenotypic
CC manifestations of perturbing the normal function of NHP in the body.
CC Sequences AAF83635-647 represent NHP nucleic acid sequences.
XX
SQ Sequence 2700 BP; 640 A; 700 C; 731 G; 629 T; 0 other;

Query Match 45.2%; Score 2199.4; DB 22; Length 2700;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 2699; Conservative 0; Mismatches 1; Indels 489; Gaps 1;

QY 1308 ATGACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCAGCGCATC 1367
DB 1 ATGACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCAGCGCATC 60

QY 1368 CGTGTGACCAGTGAAGTTCACACTGCATTAAGCTGATTAATAATGTACACATGGGAGAAA 1427
DB 61 CGTGTGACCAGTGAAGTTCACACTGCATTAAGCTGATTAATAATGTACACATGGGAGAAA 120

QY 1428 CCATTTGCAAAAATCATTTGAAGACCTTAAGAAGGAAGAAAGGAAACTATTGGAGAAGTGC 1487
DB 121 CCATTTGCAAAAATCATTTGAAGACCTTAAGAAGGAAGAAAGGAAAGCTATTGGAGAAGTGC 180

QY 1488 GGGCTTGTCAGAGCCTGACAAGTATAACCTTGTTTCATCATCCCCACAGTGGCCACAGCG 1547
DB 181 GGGCTTGTCAGAGCCTGACAAGTATAACCTTGTTTCATCATCCCCACAGTGGCCACAGCG 240

QY 1548 GTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGC 1607
DB 241 GTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGC 300

QY 1608 ATGCTGGCCCTCCTTGAATCTCCTTCGGCTGTCAGTGTCTTTTGTGCCTATTGCAGTCAA 1667
DB 301 ATGCTGGCCCTCCTTGAATCTCCTTCGGCTGTCAGTGTCTTTTGTGCCTATTGCAGTCAA 360

QY 1668 GGTCTCACGAATTCCAAGTCTGCAGTGATGAGGTTCAAGAAGTTTTTTCCTCCAGGAGAGC 1727
DB 361 GGTCTCACGAATTCCAAGTCTGCAGTGATGAGGTTCAAGAAGTTTTTTCCTCCAGGAGAGC 420

QY 1728 CCTGTTTCTATGTCCAGACATTACAAGACCCACAGAAAGCTCTGGTCTTTTGAGGAGGCC 1787
DB 421 CCTGTTTCTATGTCCAGACATTACAAGACCCACAGAAAGCTCTGGTCTTTTGAGGAGGCC 480

QY 1788 ACCTTGTCATGGCAACAGACCTGTCCCGGGATCGTCAATGGGGCACTGGAGCTGGAGAGG 1847
DB 481 ACCTTGTCATGGCAACAGACCTGTCCCGGGATCGTCAATGGGGCACTGGAGCTGGAGAGG 540

QY 1848 AACGGGCATGCTTCTGAGGGGATGACCAGGCCTAGAGATGCCCTCGGGCCAGAGGAAGAA 1907
DB 541 AACGGGCATGCTTCTGAGGGGATGACCAGGCCTAGAGATGCCCTCGGGCCAGAGGAAGAA 600

QY 1908 GGGAAACAGCCTGGGCCACAGATTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATG 1967
DB 601 GGGAAACAGCCTGGGCCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATG 660

QY 1968 TTAGGGGTCTCGGGCAACACGGGGAGTGGTAAGAGCAGCCTGTTGTACGCCATCCTGGAG 2027
DB 661 TTAGGGGTCTCGGGCAACACGGGGAGTGGTAAGAGCAGCCTGTTGTACGCCATCCTGGAG 720

QY 2028 GAGATGCACCTTGCTCGAGGGTCGGTGGGGTGCAGGAAGCCTGGCCTATGTCCCCCAG 2087
DB 721 GAGATGCACCTTGCTCGAGGGCTCGGTGGGGTGCAGGAAGCCTGGCCTATGTCCCCCAG 780

QY 2088 CAGGCCTGGATCGTCAGCGGGAACATCAGGGGAGAACAATCCTCATGGAGGCGCATATGAC 2147
DB 781 CAGGCCTGGATCGTCAGCGGGAACATCAGGGGAGAACAATCCTCATGGAGGCGCATATGAC 840

QY 2148 AAGGCCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACCTCTG 2207
DB 841 AAGGCCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACCTCTG 900

QY 2208 CCCTTTGGAGACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGGCAGAAA 2267
DB 901 CCCTTTGGAGACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGGCAGAAA 960

QY 2268 CAGAGGATCAGCCTGGCCGCGCGCTCTATTCCGACCGTCAGATCTACCTGCTGGAGCAG 2327
DB 961 CAGAGGATCAGCCTGGCCGCGCGCTCTATTCCGACCGTCAGATCTACCTGCTGGAGCAG 1020

QY 2328 CCCCTGTCTGTGTGGACGCCACAGTGGGGGAAGCACATTTTGTGAGGAGTGCAJTAAGAAG 2387
DB 1021 CCCCTGTCTGTGTGGACGCCACAGTGGGGGAAGCACATTTTGTGAGGAGTGCAJTAAGAAG 1080

QY 2388 ACACTCAGGGGGAAGACGGTCGTCTGCTGGTGACCCACCAAGTGCAGTACTTAGAATTTTGT 2447
DB 1081 ACACTCAGGGGGAAGACGGTCGTCTGCTGGTGACCCACCAAGTGCAGTACTTAGAATTTTGT 1140

QY 2448 GGCCAGATCATTTTGTGAAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTA 2507
DB 1141 GGCCAGATCATTTTGTGAAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTA 1200

QY 2508 ATGCAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAAGATGCACAAGGAAGCCACTTCG 2567
DB 1201 ATGCAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAAGATGCACAAGGAAGCCACTTC - 1259

QY 2568 GACATGTTGCAGGACACAGCAAAAGATAGCAGAGAGCCAAAGGTAGAAAAGTCAGGCTCTG 2627
DB 1260 ----- 1259

QY 2628 GCCACCTCCCTGGAAGAGTCTCTCAACGGGAAATGCTGTGCCGGAGCATCAGCTCACACAG 2687
DB 1260 ----- 1259

QY 2688 GAGGAGGAGATGGAAGAAGGCTCCTTGAGTTGGAGGGTCTACCACCACACATCCAGGCA 2747
DB 1260 ----- 1259

QY 2748 GCTGGAGGTTACATGGTCTCTTGCAATAATTTTCTCTTCGTGGTGCTGATCGTCTCTTA 2807
DB 1260 ----- 1259

QY 2808 ACGATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGC 2867
DB 1260 ----- 1259

QY 2868 AGCCGAGAGAGCAATGGAAACCATGGCAGACCTGGGCAACATTCAGACAATCCTCAACTG 2927
DB 1260 ----- 1259

QY 2928 TCCTTCTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGC 2987
DB 1260 ----- 1259

QY 2988 TCCTCAGGGATTTTCACCAAAGTCACGAGGAAGGCATCCACGGCCCTGCACAAACAAGCTC 3047
DB 1260 ----- 1259

QY 3048 TTCAACAAGGTTTTCCGCTGCCCCCATGAGTTTCTTTTGACACCATCCCAATAGGCCGGCTT 3107

Db 1260 -----GTTTTCCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCCGGCTT 1311
QY 3108 TTGAAC TGCTTCGCAGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAG 3167
Db 1312 TTGAAC TGCTTCGCAGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAG 1371
QY 3168 CAGTTCCCTGGTCCTGTCTTAATGGTGATCGCCGTCCTGTTGATGTCAAGTGCTGTCT 3227
Db 1372 CAGTTCCCTGGTCCTGTCTTAATGGTGATCGCCGTCCTGTTGATGTCAAGTGCTGTCT 1431
QY 3228 CCATATATCCTGTTAATGGAGGCCATAATCATGGTTATTTGCTTCATTTATTATATGATG 3287
Db 1432 CCATATATCCTGTTAATGGAGGCCATAATCATGGTTATTTGCTTCATTTATTATATGATG 1491
QY 3288 TTCAAGAAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTTC 3347
Db 1492 TTCAAGAAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTTC 1551
QY 3348 TCCCACATCCTCAATTCTCTGCAAGGCCTGAGCTCCATCCATGCTATGGAAAACTGAA 3407
Db 1552 TCCCACATCCTCAATTCTCTGCAAGGCCTGAGCTCCATCCATGCTATGGAAAACTGAA 1611
QY 3408 GACTTTCATCAGCCAGTTTAAAGAGGCTGACTGATCGCCAGAATACTACCTGCTGTTGTTT 3467
Db 1612 GACTTTCATCAGCCAGTTTAAAGAGGCTGACTGATCGCCAGAATACTACCTGCTGTTGTTT 1671
QY 3468 CTATCTTCCACACGATGGATGGCATTGAGGCTGGAGATCATGACCAACCTTGTGACCTTG 3527
Db 1672 CTATCTTCCACACGATGGATGGCATTGAGGCTGGAGATCATGACCAACCTTGTGACCTTG 1731
QY 3528 GCTGTTGCCCTGTTCTGTTGGCTTTTGGCATTTCTCCACCCCTTACTCCTTTAAAGTCATG 3587
Db 1732 GCTGTTGCCCTGTTCTGTTGGCTTTTGGCATTTCTCCACCCCTTACTCCTTTAAAGTCATG 1791
QY 3588 GCTGTCAACATCGTGCTGCAGCTGGCTGCCAGCTTCCAGGCCACTGCCCGGATTGGCTTG 3647
Db 1792 GCTGTCAACATCGTGCTGCAGCTGGCTGCCAGCTTCCAGGCCACTGCCCGGATTGGCTTG 1851
QY 3648 GAGACAGAGSCACAGTTTACGGCTGTAGAGAGGATACTGCAGTACATGAAGATGTGTGC 3707
Db 1852 GAGACAGAGSCACAGTTTACGGCTGTAGAGAGGATACTGCAGTACATGAAGATGTGTGC 1911
QY 3708 TCGGAAGCTCCTTTACACATGGAAGSCACAAGTTGTCCCCAGGGGTGCCACAGCATGG 3767
Db 1912 TCGGAAGCTCCTTTACACATGGAAGSCACAAGTTGTCCCCAGGGGTGCCACAGCATGG 1971
QY 3768 GAAATCATATTTCAGGATTATCACAATGAAATACAGAGACAACACACCCACCGTCTTCAC 3827
Db 1972 GAAATCATATTTCAGGATTATCACAATGAAATACAGAGACAACACACCCACCGTCTTCAC 2031
QY 3828 GGCATCAACCTGACCATCCGGCGGCCACCAAGTGGTGGGCATCTGTTGGGAAGGACGGGCTCT 3887
Db 2032 GGCATCAACCTGACCATCCGGCGGCCACCAAGTGGTGGGCATCTGTTGGGAAGGACGGGCTCT 2091
QY 3888 GGAAGTCTCCTTGGGCATGGCTCTCTTCCGCCTGGTGGAGCCCATGGCAGGCCGGATT 3947
Db 2092 GGAAGTCTCCTTGGGCATGGCTCTCTTCCGCCTGGTGGAGCCCATGGCAGGCCGGATT 2151
QY 3948 CTCATTGAGCGCGTGGACATTTGCAGCATCGGCCTGGAGGACTTGGGTCCCAAGCTCTCA 4007
Db 2152 CTCATTGAGCGCGTGGACATTTGCAGCATCGGCCTGGAGGACTTGGGTCCCAAGCTCTCA 2211
QY 4008 GTGATCCTCAAGATCCAGTGCTGCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTT 4067
Db 2212 GTGATCCTCAAGATCCAGTGCTGCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTT 2271
QY 4068 GACCGTCACTGACCCAGCAGATCTGGGATGCCTTGGAGAGGACATTCCTGACCAAGGCC 4127
Db 2272 GACCGTCACTGACCCAGCAGATCTGGGATGCCTTGGAGAGGACATTCCTGACCAAGGCC 2331
QY 4128 ATCTCAAGTTCCTCCCAAAAGCTGCATACAGATGTGGTGGAAACGGTGGAAACTTCTCT 4187

Db 2332 ATCTCAAAGTTCCCAAAAAGCTGCATACAGATGTGGTGAAAAACGGTGGAAACTTCTCT 2391
QY 4188 GTGGGGAGAGGACAGCTGCTCTGCATTTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATC 4247
Db 2392 GTGGGGAGAGGACAGCTGCTCTGCATTTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATC 2451
QY 4248 CTTATCGATGAAGCCACACAGCCTCCATTTGACATGGAGACAGACACCCCTGATCCAGCGCAC 4307
Db 2452 CTTATCGATGAAGCCACACAGCCTCCATTTGACATGGAGACAGACACCCCTGATCCAGCGCAC 2511
QY 4308 ATCCGTGAAGCCTTCCAGGGCTGCACCGTCTCGTCATTTGCCACCGTGTCAACCACTGTG 4367
Db 2512 ATCCGTGAAGCCTTCCAGGGCTGCACCGTCTCGTCATTTGCCACCGTGTCAACCACTGTG 2571
QY 4368 CTGAAC TGTGACCCACATCCTGGTTATTTGGCAATGGGAAGGTGGTAGAATTTGATCGGCCG 4427
Db 2572 CTGAAC TGTGACCCACATCCTGGTTATTTGGCAATGGGAAGGTGGTAGAATTTGATCGGCCG 2631
QY 4428 GAGTACTGCGGAAGAAGCCTGGGTCTATTTGGCAATGGGAAGGTGGTAGAATTTGATCGGCCG 4487
Db 2632 GAGTACTGCGGAAGAAGCCTGGGTCTATTTGGCAATGGGAAGGTGGTAGAATTTGATCGGCCG 2691
QY 4488 TCACTGAGA 4496
Db 2692 TCACTGAGA 2700
RESULT 15
AAF83636
ID AAF83636 standard; cDNA; 2115 BP.
XX
AC AAF83636;
XX
DT 23-JUL-2001 (first entry)
XX
DE Novel human transporter protein (NHP) encoding cDNA.
XX
KW Novel human protein; transporter protein; NHP; therapeutic; diagnostic;
KW gene therapy; ss.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..2115
ET /*tag= a
XX
PN WO200132706-A2.
XX
PD 10-MAY-2001.
XX
PF 31-OCT-2000; 2000WO-US29852.
XX
PR 02-NOV-1999; 99US-0163018.
XX
PA (LEXI-) LEXICON GENETICS INC.
PI Harras M, Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B;
PI Sands AT;
XX
DR WPI; 2001-343477/36.
DR P-PSDB; AAB62457.
XX
PT Novel isolated human polynucleotide sequences encoding polypeptides
PT that share sequence similarity with mammalian multidrug resistance
PT proteins and cellular transporter proteins, useful as probe or primer
PT
XX
PS Disclosure; Page 27-28; 59pp; English.
XX
CC The invention relates to novel human transporter proteins (NHP) and
CC polynucleotides encoding the NHPs, fragments of NHP nucleotide sequences
CC can be used in conjunction with PCR to screen libraries, isolate clones
CC and prepare cloning and sequencing templates. The NHP oligonucleotides

CC	can also be used as hybridization probes for screening libraries, for									
CC	assessing gene patterns and for preparing antisense nucleic acid									
CC	molecules. The NHP nucleotide sequences are also useful in screening									
CC	techniques for drugs which treats symptomatic or phenotypic									
CC	manifestations of perturbing the normal function of NHP in the body.									
CC	Sequences AAF83635-647 represent NHP nucleic acid sequences.									
XX										
SQ	Sequence 2115 BP; 518 A; 532 C; 578 G; 487 T; 0 other;									
	Query Match	43.4%;	Score 2111.8;	DB 22;	Length 2115;					
	Best Local Similarity	99.9%;	Pred. No. 0;							
	Matches 2113;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;					
QY	1308	ATGACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGACGACCGCATC	1367							
Db	1	ATGACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGACGACCGCATC	60							
QY	1368	CGTGTGACCAGTGAAGTTCTCACTTGCAATTAAGCTGATTAAAAATGTACACATGGGAGAAA	1427							
Db	61	CGTGTGACCAGTGAAGTTCTCACTTGCAATTAAGCTGATTAAAAATGTACACATGGGAGAAA	120							
QY	1428	CCATTGCAAAAATCATTGAAGACCTAAGAAAGGAAGGAAAGAACTATTGGAGAAGTGC	1487							
Db	121	CCATTGCAAAAATCATTGAAGACCTAAGAAAGGAAGGAAAGAACTATTGGAGAAGTGC	180							
QY	1488	GGGCTTGTCCAGAGCCCTGACAAGTATAACCTTGTTTCATCATCCCCACAGTGGCCACAGCG	1547							
Db	181	GGGCTTGTCCAGAGCCCTGACAAGTATAACCTTGTTTCATCATCCCCACAGTGGCCACAGCG	240							
QY	1548	GTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAGTCAACAGCGTCAATGGCCTTCAGC	1607							
Db	241	GTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAGTCAACAGCGTCAATGGCCTTCAGC	300							
QY	1608	ATGCTGGCCTCCTTGAATCTCCTTCGGCTGTCAGTGTCTTTGTGCCTATTGCAGTCAAA	1667							
Db	301	ATGCTGGCCTCCTTGAATCTCCTTCGGCTGTCAGTGTCTTTGTGCCTATTGCAGTCAAA	360							
QY	1668	GGTCTCACGAATTCGAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTTTCCTCCAGGAGAGC	1727							
Db	361	GGTCTCACGAATTCGAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTTTCCTCCAGGAGAGC	420							
QY	1728	CCTGTTTCTATGTCCAGACATTAACAAGACCCACGAAAGCTCTGGTCTTTGAGGAGGCC	1787							
Db	421	CCTGTTTCTATGTCCAGACATTAACAAGACCCACGAAAGCTCTGGTCTTTGAGGAGGCC	480							
QY	1788	ACCTTGTCATGGCAACAGACCTGTCCCGGGATCGTCAATGGGGCACTGGAGCTGGAGAGG	1847							
Db	481	ACCTTGTCATGGCAACAGACCTGTCCCGGGATCGTCAATGGGGCACTGGAGCTGGAGAGG	540							
QY	1848	AACGGGCATGCTTCTGAGGGGATGACCAAGCCCTAGAGATGCCCTCGGGCCAGAGGAAGAA	1907							
Db	541	AACGGGCATGCTTCTGAGGGGATGACCAAGCCTTAGAGATGCCCTCGGGCCAGAGGAAGAA	600							
QY	1908	GGGAACAGCCTGGGCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATG	1967							
Db	601	GGGAACAGCCTGGGCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATG	660							
QY	1968	TTAGGGGTCTGCGGCAACACGGGGAGTGGTAAGAGCAGCCTGTTGTGAGCCATCCTGGAG	2027							
Db	661	TTAGGGGTCTGCGGCAACACGGGGAGTGGTAAGAGCAGCCTGTTGTGAGCCATCCTGGAG	720							
QY	2028	GAGATGCACTTGCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCTGGCCCTATGTCCCCCAG	2087							
Db	721	GAGATGCACTTGCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCTGGCCCTATGTCCCCCAG	780							
QY	2088	CAGGCCTGGATCGTCAGCGGGAACATCAGGGAGAACATCCTCATGGGAGGGCGCATATGAC	2147							
Db	781	CAGGCCTGGATCGTCAGCGGGAACATCAGGGAGAACATCCTCATGGGAGGGCGCATATGAC	840							
QY	2148	AAGGCCCGATACCTCCAGGTGCTCCACTGCTGCCTCCCTGAATCGGGACCTGGAACCTCTG	2207							
Db	841	AAGGCCCGATACCTCCAGGTGCTCCACTGCTGCCTCCCTGAATCGGGACCTGGAACCTCTG	900							

Qy	2208	CCCTTTGGAGACATGACAGAGATTGGAGAGCGGGCCCTCAACCTCTCTGGGGGGCAGAAA	2267
Db	901	CCCTTTGGAGACATGACAGAGATTGGAGAGCGGGCCCTCAACCTCTCTGGGGGGCAGAAA	960
Qy	2268	CAGAGGATCAGCCTGGCCGCGCGTCTATTCCGACCGTCAGATCTACCTGCTGGACGAC	2327
Db	961	CAGAGGATCAGCCTGGCCGCGCGTCTATTCCGACCGTCAGATCTACCTGCTGGACGAC	1020
Qy	2328	CCCCTGTCTGCTGTGGACGCCACAGTGGGGAAAGCACATTTTGGAGAGTGCATTAAAGAG	2387
Db	1021	CCCCTGTCTGCTGTGGACGCCACAGTGGGGAAAGCACATTTTGGAGAGTGCATTAAAGAG	1080
Qy	2388	ACACTCAGGGGAAGACCGTCTGCTGGTGACCCACCACTGCAGTACTTAGAATTTTGT	2447
Db	1081	ACACTCAGGGGAAGACCGTCTGCTGGTGACCCACCACTGCAGTACTTAGAATTTTGT	1140
Qy	2448	GGCCAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTA	2507
Db	1141	GGCCAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTA	1200
Qy	2508	ATGCAGAAAAAGGGGAATATGCCAACTTATCCAGAAGATGCACAAAGGAAAGCCACTCG	2567
Db	1201	ATGCAGAAAAAGGGGAATATGCCAACTTATCCAGAAGATGCACAAAGGAAAGCCACTCG	1260
Qy	2568	GACATGTTGCAGGACACAGCAAAAGATAGCAGAGAAGCAAAAAGGTAGAAAGTCAGGCTCTG	2627
Db	1261	GACATGTTGCAGGACACAGCAAAAGATAGCAGAGAAGCAAAAAGGTAGAAAGTCAGGCTCTG	1320
Qy	2628	GCCACCTCCCTGGAAAGTCTCTCAACGGGAAATGCTGTGCCGGAGCATCAGCTCACACAG	2687
Db	1321	GCCACCTCCCTGGAAAGTCTCTCAACGGGAAATGCTGTGCCGGAGCATCAGCTCACACAG	1380
Qy	2688	GAGGAGGAGATGGAAGAGGCTCCTTGAGTTGGAGGTCTACCACCACCTACATCCAGGCA	2747
Db	1381	GAGGAGGAGATGGAAGAGGCTCCTTGAGTTGGAGGTCTACCACCACCTACATCCAGGCA	1440
Qy	2748	GCTGGAGGTTACATGCTCTTTCATATAATTTTCTCTTCGTGGTGTGATCGTCTTTTA	2807
Db	1441	GCTGGAGGTTACATGCTCTTTCATATAATTTTCTCTTCGTGGTGTGATCGTCTTTTA	1500
Qy	2808	ACGATCTTCAGCTTCCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGACCAATAGC	2867
Db	1501	ACGATCTTCAGCTTCCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGACCAATAGC	1560
Qy	2868	AGCCGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCAGACAATCCTCAACTG	2927
Db	1561	AGCCGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCAGACAATCCTCAACTG	1620
Qy	2928	TCCTTCTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGC	2987
Db	1621	TCCTTCTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGC	1680
Qy	2988	TCCTCAGGGATTTTTCACCAAGTCACGAGGAAGGCATCCACGGCCCTGCACACAAGCTC	3047
Db	1681	TCCTCAGGGATTTTTCACCAAGTCACGAGGAAGGCATCCACGGCCCTGCACACAAGCTC	1740
Qy	3048	TTCAACAAGGTTTTCGGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCCGGCTT	3107
Db	1741	TTCAACAAGGTTTTCGGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCCGGCTT	1800
Qy	3108	TTGAACTGCTTCGCAAGGGACTTGGAAACAGCTGGACAGCTCTTGCCCCATCTTTTCAGAG	3167
Db	1801	TTGAACTGCTTCGCAAGGGACTTGGAAACAGCTGGACAGCTCTTGCCCCATCTTTTCAGAG	1860
Qy	3168	CAGTTCCTGGTCCCTTAATGGTGTATCGCCCGTCCCTGTGATTGTGTCAGTGTCTGTCT	3227
Db	1861	CAGTTCCTGGTCCCTTAATGGTGTATCGCCCGTCCCTGTGATTGTGTCAGTGTCTGTCT	1920
Qy	3228	CCATATATCCTGTTAATGGAGCCATAATCATGGTTATTGCTTCATTTTATATATGATG	3287
Db	1921	CCATATATCCTGTTAATGGAGCCATAATCATGGTTATTGCTTCATTTTATATATGATG	1980
Qy	3288	TTCAAGAAGGCCATCGGTGTGTTCAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTTC	3347

Db	1981	TTCAAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTC	2040
QY	3348	TCCCACATCCTCAATTCTCTGCAAGGCCCTGAGCTCCATCCATGTCTATGGAAAACTGAA	3407
Db	2041	TCCCACATCCTCAATTCTCTGCAAGGCCCTGAGCTCCATCCATGTCTATGGAAAACTGAA	2100
QY	3408	GACTTCATCAGCCAG	3422
Db	2101	GACTTCATCAGCCAG	2115

Search completed: July 24, 2003, 06:10:47
Job time : 969 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2003, 05:14:11 ; Search time 12054 Seconds
(without alignments)
11738.663 Million cell updates/sec

Title: US-10-087-782A-1
Perfect score: 4862
Sequence: 1 actgggataaagcaaga.....tataaaactaaggaaaaactc 4862

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*				GenEmbl:*			
1:	gb_ba:*	1:	gb_ba:*	21:	gb_ro:*	21:	gb_ro:*
2:	gb_htg:*	2:	gb_htg:*	22:	gb_sts:*	22:	gb_sts:*
3:	gb_in:*	3:	gb_in:*	23:	gb_sy:*	23:	gb_sy:*
4:	gb_om:*	4:	gb_om:*	24:	gb_un:*	24:	gb_un:*
5:	gb_ov:*	5:	gb_ov:*	25:	gb_vi:*	25:	gb_vi:*
6:	gb_pat:*	6:	gb_pat:*	26:	em_ba:*	26:	em_ba:*
7:	gb_ph:*	7:	gb_ph:*	27:	em_fun:*	27:	em_fun:*
8:	gb_pl:*	8:	gb_pl:*	28:	em_hum:*	28:	em_hum:*
9:	gb_pr:*	9:	gb_pr:*	29:	em_in:*	29:	em_in:*
10:	gb_ro:*	10:	gb_ro:*	30:	em_mu:*	30:	em_mu:*
11:	gb_sts:*	11:	gb_sts:*	31:	em_om:*	31:	em_om:*
12:	gb_sy:*	12:	gb_sy:*	32:	em_or:*	32:	em_or:*
13:	gb_un:*	13:	gb_un:*	33:	em_ov:*	33:	em_ov:*
14:	gb_vi:*	14:	gb_vi:*	34:	em_pat:*	34:	em_pat:*
15:	em_ba:*	15:	em_ba:*	35:	em_ph:*	35:	em_ph:*
16:	em_fun:*	16:	em_fun:*	36:	em_pl:*	36:	em_pl:*
17:	em_hum:*	17:	em_hum:*	37:	em_ro:*	37:	em_ro:*
18:	em_in:*	18:	em_in:*	38:	em_sts:*	38:	em_sts:*
19:	em_mu:*	19:	em_mu:*	39:	em_un:*	39:	em_un:*
20:	em_om:*	20:	em_om:*	40:	em_vi:*	40:	em_vi:*
21:	em_or:*	21:	em_or:*	41:	em_htg_hum:*	41:	em_htg_hum:*
22:	em_ov:*	22:	em_ov:*	42:	em_htg_inv:*	42:	em_htg_inv:*
23:	em_pat:*	23:	em_pat:*	43:	em_htg_other:*	43:	em_htg_other:*
24:	em_ph:*	24:	em_ph:*	44:	em_htg_mus:*	44:	em_htg_mus:*
25:	em_pl:*	25:	em_pl:*	45:	em_htg_pln:*	45:	em_htg_pln:*
26:	em_ro:*	26:	em_ro:*	46:	em_htg_rod:*	46:	em_htg_rod:*
27:	em_sts:*	27:	em_sts:*	47:	em_htg_mam:*	47:	em_htg_mam:*
28:	em_un:*	28:	em_un:*	48:	em_htg_vrt:*	48:	em_htg_vrt:*
29:	em_vi:*	29:	em_vi:*	49:	em_sy:*	49:	em_sy:*
30:	em_htg_hum:*	30:	em_htg_hum:*	50:	em_htgo_hum:*	50:	em_htgo_hum:*
31:	em_htg_inv:*	31:	em_htg_inv:*	51:	em_htgo_mus:*	51:	em_htgo_mus:*
32:	em_htg_other:*	32:	em_htg_other:*	52:	em_htgo_other:*	52:	em_htgo_other:*
33:	em_htg_mus:*	33:	em_htg_mus:*				
34:	em_htg_pln:*	34:	em_htg_pln:*				
35:	em_htg_rod:*	35:	em_htg_rod:*				
36:	em_htg_mam:*	36:	em_htg_mam:*				
37:	em_htg_vrt:*	37:	em_htg_vrt:*				
38:	em_sy:*	38:	em_sy:*				
39:	em_htgo_hum:*	39:	em_htgo_hum:*				
40:	em_htgo_mus:*	40:	em_htgo_mus:*				
41:	em_htgo_other:*	41:	em_htgo_other:*				

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	4862	100.0	4862	9	AY040219	AY040219 Homo sapi
2	4523.4	93.0	4590	9	AF367202	AF367202 Homo sapi
3	4392.8	90.3	4492	9	AF352582	AF352582 Homo sapi
4	4391.4	90.3	4638	6	AX468400	AX468400 Sequence
5	4285.4	88.1	4476	9	AF411579	AF411579 Homo sapi
6	4142.6	85.2	4149	6	AX135184	AX135184 Sequence
7	4133.6	85.0	5020	6	AX135192	AX135192 Sequence
8	3964.2	81.5	4074	6	AX411822	AX411822 Sequence
9	3836.4	78.9	4083	6	AX468402	AX468402 Sequence
10	3185.8	65.5	3189	6	AX135172	AX135172 Sequence
11	3156.2	64.9	3660	6	AX135190	AX135190 Sequence
12	3066.6	63.1	3075	6	AX135182	AX135182 Sequence
13	2927.6	60.2	2937	6	AX135180	AX135180 Sequence
14	2213.2	45.5	2448	6	AX135186	AX135186 Sequence
15	2213.2	45.5	2586	6	AX135188	AX135188 Sequence
16	2199.4	45.2	2700	6	AX135178	AX135178 Sequence
17	2111.8	43.4	2115	6	AX135170	AX135170 Sequence
18	2102.2	43.2	2699	6	AX392930	AX392930 Sequence
19	2010	41.3	2010	6	AX411837	AX411837 Sequence
20	1973.8	40.6	1977	6	AX135168	AX135168 Sequence
21	1716	35.3	1824	9	HSM801041	AL117406 Homo sapi
22	1279.2	26.3	4101	6	AX210135	AX210135 Sequence
23	1273.2	26.2	4127	9	AF411577	AF411577 Homo sapi
24	1259.4	25.9	1488	6	AX135174	AX135174 Sequence
25	1259.4	25.9	1626	6	AX135176	AX135176 Sequence
26	1218.2	25.1	5168	9	AY040220	AY040220 Homo sapi
27	1143.6	23.5	4048	9	AF411578	AF411578 Homo sapi
28	1132.8	23.3	4008	6	AX210141	AX210141 Sequence
29	1128.4	23.2	4034	9	AF395908	AF395908 Homo sapi
30	1063	21.9	3660	6	AX210137	AX210137 Sequence
31	1057.4	21.7	3753	6	AX210139	AX210139 Sequence
32	936	19.3	3886	9	AF395909	AF395909 Homo sapi
33	803.4	16.5	5400	10	AB019003	AB019003 Mus muscu
34	801	16.5	5744	10	AB020209	AB020209 Rattus no
35	768	15.8	4781	6	AR091520	AR091520 Sequence
36	768	15.8	4781	6	AR136853	AR136853 Sequence
37	768	15.8	4847	6	AR099619	AR099619 Sequence
38	766.4	15.8	5826	9	HSU83661	U83661 Homo sapien
39	766.4	15.8	5881	9	AF146074	AF146074 Homo sapi
40	764.8	15.7	5728	9	AB019002	AB019002 Homo sapi
41	764.8	15.7	5838	9	AF104942	AF104942 Homo sapi
42	722.2	14.9	4939	9	AB005659	AB005659 Homo sapi
43	641.2	13.2	732	6	AX411832	AX411832 Sequence
44	629.8	13.0	1698	6	AX210143	AX210143 Sequence
45	624	12.8	673	6	AX411828	AX411828 Sequence

ALIGNMENTS

RESULT 1	AY040219	AY040219	4862 bp	mrna	linear	PRI 01-AUG-2001
AY040219						
LOCUS						
DEFINITION	Homo sapiens	ATP-binding cassette transporter sub-family C member				
	11 (ABCC11)	mrna, complete cds.				
ACCESSION	AY040219					
VERSION	AY040219.1	GI:15027828				
KEYWORDS						
SOURCE						
ORGANISM	Homo sapiens.					
	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 4862)					
AUTHORS	Tammur, J., Prades, C., Arnould, I., Rzhetsky, A., Hutchinson, A.,					
	Adachi, M., Schuetz, J.D., Swoboda, K.J., Ptacek, L.J., Rosier, M.,					

TITLE	Dean,M. and Allikmets,R. Two new genes from the human ATP-binding cassette transporter superfamily, ABCC11 and ABCC12, tandemly duplicated on chromosome 16q12
JOURNAL	Gene 273 (1), 89-96 (2001)
MEDLINE	21376129
PUBMED	11483364
REFERENCE	2 (bases 1 to 4862)
AUTHORS	Tammur,J., Prades,C., Arnould,I., Rosier,M., Dean,M. and Allikmets,R.
TITLE	Direct Submission
JOURNAL	Submitted (14-JUN-2001) Ophthalmology, Columbia University, Eye Research Addition, 630 West 168th Street, New York, NY 10032, USA
FEATURES	Location/Qualifiers
source	1..4862 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="16" /map="16q12" /tissue_type="testis"
gene	1..4862 /gene="ABCC11"
CDS	351..4499 /gene="ABCC11" /note="CFTR/MRP" /codon_start=1 /product="ATP-binding cassette transporter sub-family C member 11" /protein_id="AAK76739.1" /db_xref="GI:15027829" /translation="MTRKRTYWVPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWS QERNPEAPGRAVPWPWKYDAALRTMIPKPRFPAPQPLDNAGLFSYLTVSWLTP LMIQSLRSRLDENTIPPLSVHDASDKNVQRLHRLWEEVSRRGIEKASVLLVMLRFQR TRLIFDALLGICFCIASVLGPILIIIPKILEYSEEQLGNVHVHGLCFALFLSECVKSL SFSSSWIINQRTAIRFRAAVSFAFEKLIQPKSVIHITSGEAISFFTGDNVYLFEGVC YGPLVITCASLVICSISSYFIIGYTAFTAILCYLLVFPFLAVFMTRNAVKAQHHTSEV SDQIRVTSVLTICIKLIKMYTWEKPFKIIEDLRRKERKLLKCGLVQSLTSITLFI IPTVATVAVVLIHTSLKLTASMAFSMLASLNLLRLSVFFVPIAVKGLTNSKSAVMR FKKFFLQESPVFVQTLQDPSKALVFEELTSWQOTCPGIVNGALELERNGHASEGMT RPRDALGPEEENSLGPELHKINLVVSKGMLGVCGNTGSGKSSLLSAILSEMHLLLEG SVGVQGSLAYVPOQAWIVSGNIRENILMGAYDKARYLQVLHCCSLNRDLELLPFGDM TEIGERGLNLSGQKQISLARAVYSRQIYLLDDPLSADVDAHVGKHIFEECIKKTLR GKTVVLTQHQLYLEFCGQIILLENGKICENGTHSELMQKKGKYZAQLIQMHKEATSD MLODTAKIAEKPQVESQALATSLSEESLGNNAVPEHQLTQEEEMEESLSWRVYHHYIQ AAGYMWVSCIIFFFVVLIVFLTIFSFVWLSYWEQSGSTNSSRESNGTMADLGNIADN POLSFYQLVYGLNALLLICVGVCSGGIFTKVRKASTALHNKLFNKVFCRPMSEFDTI PIGRLLNCFAGLEQLDQLLPFISEQLVLSLMVIAVLIVSVLSPYILLMGAIIMVI CFIYYMMFKKAIGVFKRLNYSRSPFLSHILNSLQGLSSIHVYKGTEDFISQFKRLTD AQNYYLLFLSSTRWMLRLEIMTNLTVLAVLFAVFGISSTPYSPFKVMAVNIIVLQLA SSFQATARIGLETEAQTFAVERILQYMKMCVSEAPLHMEGTSCPOGWPHGEIIFQDY HMKYRDNTPTVLGINLTIIRGHEVVGIVGRGTSGKSSLGMAFLRLVEPMAGRILIDGV DICSIGLEDLRKLSVIPQDPVLLSGTIRFNLDPFDRHTDQOIWDALERTFLTKAISK FPKKLHTDVVENGNFSVGERQLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTI REAFQGCITVLVIAHRVTTVLNCDHILVMGNKGVVEFDRPEVLRKKPGSLFAALMATAT SSLR"
BASE COUNT	1177 a 1211 c 1294 g 1180 t
ORIGIN	
Query Match	100.0%; Score 4862; DB 9; Length 4862;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 4862; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ACTGGGATAAGCAAGAAGACTGATTTTATGAGCAGGGGTTTGATACATCAAGAGGAGATT 60
Db	1 ACTGGGATAAGCAAGAAGACTGATTTTATGAGCAGGGGTTTGATACATCAAGAGGAGATT 60
QY	61 GCCCAGGATCAAGGTCGGGTGTTGGGGTGGGTGGGAGGGTGGTTAGAGAAGGTTTC 120
Db	61 GCCCAGGATCAAGGTCGGGTGTTGGGGTGGGTGGGAGGGTGGTTAGAGAAGGTTTC 120
QY	121 ACTAAGTGATTTGGGCTGAGGCCTGAGAAGATGTTAAAAAGAGGGGATCAAGCAGGC 180
Db	121 ACTAAGTGATTTGGGCTGAGGCCTGAGAAGATGTTAAAAAGAGGGGATCAAGCAGGC 180

QY	181 TAAGGAGAGGAAAGAGCAGGCACCCAAACCTCTGCATGGCCCCAATATGCTCCCTGCAGG 240
Db	181 TAAGGAGAGGAAAGAGCAGGCACCCAAACCTCTGCATGGCCCCAATATGCTCCCTGCAGG 240
QY	241 GTAGTCCCCCTCTTCTGGTGCTCAAGGCGAGATCTTAAGCTTCTTCTAACTCCTGCTGT 300
Db	241 GTAGTCCCCCTCTTCTGGTGCTCAAGGCGAGATCTTAAGCTTCTTCTAACTCCTGCTGT 300
QY	301 CTTTTCATATTCTCTGATTCTGGGAAACGAAGAATTGGCAGGAACCTGAAAATGACTAGGA 360
Db	301 CTTTTCATATTCTCTGATTCTGGGAAACGAAGAATTGGCAGGAACCTGAAAATGACTAGGA 360
QY	361 AGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGCATCGACATAG 420
Db	361 AGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGCATCGACATAG 420
QY	421 GCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGATGGCCCCCTGGA 480
Db	421 GCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGATGGCCCCCTGGA 480
QY	481 GTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGCAGCTGTCCCACCGTGGGGGAAGT 540
Db	481 GTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGCAGCTGTCCCACCGTGGGGGAAGT 540
QY	541 ATGATGCTGCCCTTGAGAACCATGATTCCCTTCCGTCGCCAAGCCGAGGTTTCTGCCCCCCC 600
Db	541 ATGATGCTGCCCTTGAGAACCATGATTCCCTTCCGTCGCCAAGCCGAGGTTTCTGCCCCCCC 600
QY	601 AGCCCTGGACAAATGCTGGCCTGTCTCTACCTCACCGTGCATGGCTCACCCCGCTCA 660
Db	601 AGCCCTGGACAAATGCTGGCCTGTCTCTACCTCACCGTGCATGGCTCACCCCGCTCA 660
QY	661 TGATCCAAAGCTTACGGAGTCGCTTAGATGAGAAACCATCCCTCCACTGTCAGTCCATG 720
Db	661 TGATCCAAAGCTTACGGAGTCGCTTAGATGAGAAACCATCCCTCCACTGTCAGTCCATG 720
QY	721 ATGCCTCAGACAAAAATGTCCAAAAGGCTTCACCGCCTTGGGAAGAAGAAGTCTCAAGGC 780
Db	721 ATGCCTCAGACAAAAATGTCCAAAAGGCTTCACCGCCTTGGGAAGAAGAAGTCTCAAGGC 780
QY	781 GAGGGATTGAAAAAGCTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGAACAAAGTTGA 840
Db	781 GAGGGATTGAAAAAGCTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGAACAAAGTTGA 840
QY	841 TTTTCGATGCACCTTCGGGCATCTGCTTCTGCATTGCCAGTGACTCGGGGCCAATATTGA 900
Db	841 TTTTCGATGCACCTTCGGGCATCTGCTTCTGCATTGCCAGTGACTCGGGGCCAATATTGA 900
QY	901 TTATACCAAAGATCCTGGAATATTTCAGAAAGAGAGTGGGGAATGTTGTCCATGGAGTGG 960
Db	901 TTATACCAAAGATCCTGGAATATTTCAGAAAGAGAGTGGGGAATGTTGTCCATGGAGTGG 960
QY	961 GACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTCTCCTCCAGTT 1020
Db	961 GACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTCTCCTCCAGTT 1020
QY	1021 GGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCTCCTTGGCCTTTG 1080
Db	1021 GGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCTCCTTGGCCTTTG 1080
QY	1081 AGAAGCTCATCCAAATTTAAGTCTGTAAATACACATCACCTCAGGAGAGGCCATCAGCTTCT 1140
Db	1081 AGAAGCTCATCCAAATTTAAGTCTGTAAATACACATCACCTCAGGAGAGGCCATCAGCTTCT 1140
QY	1141 TCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTAGTACTGATCA 1200
Db	1141 TCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTAGTACTGATCA 1200
QY	1201 CCTGCGCATCGCTGGTGCATCTGCAGCATTTTCTTCTACTTCTATTATGGATACACTGCAT 1260
Db	1201 CCTGCGCATCGCTGGTGCATCTGCAGCATTTTCTTCTACTTCTATTATGGATACACTGCAT 1260

Db 3421 AGTTTAAGAGCGTGACTGATCGCGCAGAAATAACTACCTGCTGTTGTTTCTATCTTCCACAC 3480

QY 3481 GATGGATGGCATTGAGGCTGGAGATCATGACCAACCTTGTGACCTGGCTGTTGCCCTGT 3540

Db 3481 GATGGATGGCATTGAGGCTGGAGATCATGACCAACCTTGTGACCTGGCTGTTGCCCTGT 3540

QY 3541 TCGTGGCTTTTGGCATTTCTCCACCCCTACTCCTTTAAAGTCATGGCTGTCAACATCG 3600

Db 3541 TCGTGGCTTTTGGCATTTCTCCACCCCTACTCCTTTAAAGTCATGGCTGTCAACATCG 3600

QY 3601 TGCTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCCCGGATTGGCTTGGAGACAGAGGCAC 3660

Db 3601 TGCTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCCCGGATTGGCTTGGAGACAGAGGCAC 3660

QY 3661 AGTTCACGCTGTAGAGAGGATACCTGCAGTACATGAAGATGTGTCTCGGAAGCTCCTT 3720

Db 3661 AGTTCACGCTGTAGAGAGGATACCTGCAGTACATGAAGATGTGTCTCGGAAGCTCCTT 3720

QY 3721 TACACATGGAAGGCACAAGTTGTCCCCAGGGGTGGCCACAGCATGGGGAAATCATATTTT 3780

Db 3721 TACACATGGAAGGCACAAGTTGTCCCCAGGGGTGGCCACAGCATGGGGAAATCATATTTT 3780

QY 3781 AGGATTATACATGAAATACAGAGACAACACCCACCTGCTTACGGCATCAACCTGA 3840

Db 3781 AGGATTATACATGAAATACAGAGACAACACCCACCTGCTTACGGCATCAACCTGA 3840

QY 3841 CCATCCGGGCCACGAAAGTGGTGGGCATCGTGGGAAGGACGGGCTCTGGGAAGTCCCT 3900

Db 3841 CCATCCGGGCCACGAAAGTGGTGGGCATCGTGGGAAGGACGGGCTCTGGGAAGTCCCT 3900

QY 3901 TGGGCATGCTCTCTTCCGCCTGGTGGAGCCCCATGGCAGGCCGGAATCTCATTTGACGGCG 3960

Db 3901 TGGGCATGCTCTCTTCCGCCTGGTGGAGCCCCATGGCAGGCCGGAATCTCATTTGACGGCG 3960

QY 3961 TGGACATTTGCAGCATCGGCCTGGAGGACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAG 4020

Db 3961 TGGACATTTGCAGCATCGGCCTGGAGGACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAG 4020

QY 4021 ATCCAGTGTCTCTCAGGAACCATCAGATTTCAACCTAGATCCCTTTGACCGTCAACATG 4080

Db 4021 ATCCAGTGTCTCTCAGGAACCATCAGATTTCAACCTAGATCCCTTTGACCGTCAACATG 4080

QY 4081 ACCAGCAGATCTGGGATGCCCTTGGAGAGGACATTTCTGACCAAGGCCATCTCAAAGTTTC 4140

Db 4081 ACCAGCAGATCTGGGATGCCCTTGGAGAGGACATTTCTGACCAAGGCCATCTCAAAGTTTC 4140

QY 4141 CCAAAAAGCTGCATACAGATGTGTGGAAAAACGGTGGAACCTTCTGTGGGGGAGAGGC 4200

Db 4141 CCAAAAAGCTGCATACAGATGTGTGGAAAAACGGTGGAACCTTCTGTGGGGGAGAGGC 4200

QY 4201 AGCTGCTGTGCATTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCCTTATCGATGAAG 4260

Db 4201 AGCTGCTGTGCATTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCCTTATCGATGAAG 4260

QY 4261 CCACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATCCCGTGAAGCCT 4320

Db 4261 CCACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATCCCGTGAAGCCT 4320

QY 4321 TCCAGGGCTGCACCGTGTCTGTCATTGCCACCGTGTCAACACTGTGCTGAACTGTGACC 4380

Db 4321 TCCAGGGCTGCACCGTGTCTGTCATTGCCACCGTGTCAACACTGTGCTGAACTGTGACC 4380

QY 4381 ACATCTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCGGAGGTACTGCGGA 4440

Db 4381 ACATCTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCGGAGGTACTGCGGA 4440

QY 4441 AGAAGCCTGGGTCAATTGTCGAGCCCTCATGGCCACAGCCACTCTTCACTGAGATAAG 4500

Db 4441 AGAAGCCTGGGTCAATTGTCGAGCCCTCATGGCCACAGCCACTCTTCACTGAGATAAG 4500

QY 4501 GAGATGTGAGACTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTTACACACAGGTGCAGCT 4560

Db 4501 GAGATGTGAGACTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTTACACACAGGTGCAGCT 4560

QY 4561 TCGAGGGCCACAGTCTCGGACCTTCTTGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGG 4620

Db 4561 TCGAGGGCCACAGTCTCGGACCTTCTTGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGG 4620

QY 4621 GTAAATGTAGGGGGGTGGGATTTGCTGGATGGAAACCCCTGGAATAGGCTACTTTGATGGC 4680

Db 4621 GTAAATGTAGGGGGGTGGGATTTGCTGGATGGAAACCCCTGGAATAGGCTACTTTGATGGC 4680

QY 4681 TCTCAAGACCTTAGAACCCCGAGAACCATCTAAGACATGGGATTCAGTGATCATGTGTTTC 4740

Db 4681 TCTCAAGACCTTAGAACCCCGAGAACCATCTAAGACATGGGATTCAGTGATCATGTGTTTC 4740

QY 4741 TCCTTTAAACTTACATGCTGAATAATTTTATAATAAGGTAAAGCTTATAGTTTCTGAT 4800

Db 4741 TCCTTTAAACTTACATGCTGAATAATTTTATAATAAGGTAAAGCTTATAGTTTCTGAT 4800

QY 4801 CTGTGTTAGAAAGTGTTCGAAATGCTGCTACTGACTTTGTGTAATAATAAACTAAGGAAAAC 4860

Db 4801 CTGTGTTAGAAAGTGTTCGAAATGCTGCTACTGACTTTGTGTAATAATAAACTAAGGAAAAC 4860

QY 4861 TC 4862

Db 4861 TC 4862

RESULT 2

AF367202

LOCUS AF367202

DEFINITION Homo sapiens ATP-binding cassette protein C11 (ABCC11) mRNA, complete cds.

ACCESSION AF367202

VERSION AF367202.1 GI:14280090

KEYWORDS

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 4590)

AUTHORS Yabuuchi,H., Shimizu,H., Takayanagi,S. and Ishikawa,T.

TITLE Multiple splicing variants of two new human ATP-binding cassette transporters, ABCC11 and ABCC12

JOURNAL Biochem. Biophys. Res. Commun. 288 (4), 933-939 (2001)

MEDLINE 21547789

PUBMED 11688999

REFERENCE 2 (bases 1 to 4590)

AUTHORS Yabuuchi,H. and Ishikawa,T.

TITLE Molecular cloning and characterization of human ABCC11 cDNA

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 4590)

AUTHORS Yabuuchi,H. and Ishikawa,T.

TITLE Direct Submission

JOURNAL Submitted (01-APR-2001) Biomolecular Engineering, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama, Kanagawa 226-8501, Japan

FEATURES

source

1..4590

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="16"

/map="16q12"

/tissue_type="liver"

1..4590

/gene="ABCC11"

79..4227

/gene="ABCC11"

/note="transporter"

/codon_start=1

/product="ATP-binding cassette protein C11"

/protein_id="AAK58869.1"

/db_xref="GI:14280091"

/translation="MTRKRTYWVPNSSGGLVNRGIDICDDMVSGLIYKTYTLQDGPWS

QOERNPEAPGRAAVPPWKYDAALRTMIPFRPKPRPAPQPLDNAGLSYLTVSWLTP

LMIQSLRSLDENTIPPLSVHDSKKNVQRLHRLWEEEVSRRIEKASVLLVMLRFOR
TRLIFDALLGICPCIASIVGPILIIIPKILEYSEEQLGNVAHVGLGFALFLSECVKSL
SFSSSWIINQRTAIRFRAAVSSFAPEKLIQFKSVIHITSGEAIISFTGDVNYLFEGVC
YGPLVLITCASLVICSISSYFIIGYTAFTATILCYLLVFLPLAVEMTRMAVKAQHHTSEV
SDQIRVTSVLTCIKLIKMYTWKPF AEI IEDLRRKERKLEKCLDVQSLSITSITLFI
IPTVATAVWVLHTSLKLKTASMAFSLASLNLRLSLVFFVPIAVKGLTNSKSAVMR
FKFFFLQESPVYVQTLQDPSKALVFEEATLSWQQTCPGIVNGALELERNGHASEGMT
RPRDALGPEEEGNSLGP ELHKINLVSKGMLGVCGNTGSGKSSLLSAIL EEMHLLEG
SVGVQSLAYVPOQAWIVSGNIRENILMGAYDKARYLQVLHCCSLNRDLELLPFGDM
TEIGERGLNLSGGQKORISLARAVYSDRQIYLLDDPLSAVDADAHVGKHIFEECIKKTLR
GKTVVLVTHQLOYLEFCGOIILLENGKICENGTHSELMOKKKYAQOLIKMHKEATS D
MLQDTAKIAEKPVESQALATSLESLNGNAVPEHQLTQEEEMEESLSWRVYHHYIQ
AAGGYMVCIIIPFFVVLIVFLTIFSEFWLSYWLEQSGTNSRESNGTMDLGNIADN
PQLSFYQLVYGLNALLICVGVCSGGIFTKVTRKASTALHNKLFNKVFRCPMSFFDTI
PIGRLLNCFAGDLEQLDQLLP I FSEQFLVLSLMVIAVLLIVSVLSPLLMGAIIMVI
CFIYYMMFKKAIGVFKRLENYSRSPLSFHLNSLQGLSSIHYVGKTEDFISQFKRLTD
AQNNYLLFLSSTRWMALRLEIMTNLVTLAVALFVAFGISSTPYSFVMVAVNIVLOLA
SSFQATARIGLETEAQFTAVERIILQYMKMCVSEAPLHMEGTSCPOQWPQHGEIIFQDY
HMKYRDNTPTVLHGINLTI RGHVVVGIVGR TSGSKSSLGMA LFR LVEPMAGRILIDGV
DICSIGLEDLRSKLSVIPQDPVLLSGTIRFNLDPPFDHTDQIWDALERTFLT KAI SK
FPKKLHTDVVNGGNFSVGERQLLICIA RAVLRNSKIILIDEATASIDMETDTLIQRTI
REAFQGCTVLVIAHRVTTVLNCDHILVMGNKVVEFDRPEVLRKRPGLFAALMATAT
SSLR"

BASE COUNT 1109 a 1160 c 1209 g 1112 t
ORIGIN

Query Match 93.0%; Score 4523.4; DB 9; Length 4590;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 4527; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 330 AAGAAATTGGCAGGAAC TGA AAAATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCT 389

Db 58 AAGAAATTGGCAGGAAC TGA AAAATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCT 117

Qy 390 GGTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTTAT 449

Db 118 GGTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTTAT 177

Qy 450 AAAACCTATACTCTCCAAGATGGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCA 509

Db 178 AAAACCTATACTCTCCAAGATGGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCA 237

Qy 510 GGGAGGGCAGCTGTCCCACCGTGGGGGAAGTATGATGCTGCCCTTGAGAACCATGATTCCTC 569

Db 238 GGGAGGGCAGCTGTCCCACCGTGGGGGAAGTATGATGCTGCCCTTGAGAACCATGATTCCTC 297

Qy 570 TTCCGTCCTCAAGCCGAGGTTTCTGCCCCCCAGCCCTGGACAAATGCTGGCCTGTCTCTCC 629

Db 298 TTCCGTCCTCAAGCCGAGGTTTCTGCCCCCCAGCCCTGGACAAATGCTGGCCTGTCTCTCC 357

Qy 630 TACCTCACCGTGTATGGGTCACCCCGCTCATGATGCCAAAAGCTTACGGAGTCGCTTAGAT 689

Db 358 TACCTCACCGTGTATGGGTCACCCCGCTCATGATGCCAAAAGCTTACGGAGTCGCTTAGAT 417

Qy 690 GAGAACACCATCCCTCCACTGTCAGTCCATGATGCCCTCAGACAAAAATGTCCAAGGCTT 749

Db 418 GAGAACACCATCCCTCCACTGTCAGTCCATGATGCCCTCAGACAAAAATGTCCAAGGCTT 477

Qy 750 CACCGCCTTTGGGAAGAGAAAGTCTCAAGGGCGAGGGATTGAAAAGCTTCAGTGCCTCTG 809

Db 478 CACCGCCTTTGGGAAGAGAAAGTCTCAAGGGCGAGGGATTGAAAAGCTTCAGTGCCTCTG 537

Qy 810 GTGATGCTGAGGTTCCAGAGAACAAAGGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTC 869

Db 538 GTGATGCTGAGGTTCCAGAGAACAAAGGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTC 597

Qy 870 TGCATTGCCAGTGTACTCGGGCCCAATATTGATTATACCAAGATCCTGGAATATTCAGAA 929

Db 598 TGCATTGCCAGTGTACTCGGGCCCAATATTGATTATACCAAGATCCTGGAATATTCAGAA 657

Qy 930 GACCAGTTGGGGAATGTTGTCCATGGAGTGGGACTCTGCTTTTGGCCCTTTTCTCTCCGAA 989

Db 658 GACCAGTTGGGGAATGCTGTCCATGGAGTGGGACTCTGCTTTTGGCCCTTTTCTCTCCGAA 717

Qy 990 TGTGTGAAGTCTCTGAGTTTCTCTCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGG 1049

Db 718 TGCGTGAAGTCTCTGAGTTTCTCTCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGG 777

Qy 1050 TTCCGAGCAGCTGTTTTCCTCCTTTTGCCTTTTGAGAGAGCTCATCCAAATTTAAGTCIGTAATA 1109

Db 778 TTCCGAGCAGCTGTTTTCCTCCTTTTGCCTTTTGAGAGAGCTCATCCAAATTTAAGTCIGTAATA 837

Qy 1110 CACATCACCTCAGGAGAGGCCCATCAGCTTCTTCAACGGTGATGTAAACTACCTGTGTTGAA 1169

Db 838 CACATCACCTCAGGAGAGGCCCATCAGCTTCTTCAACGGTGATGTAAACTACCTGTGTTGAA 897

Qy 1170 GGGGTGTGCTATGAGACCCCTAGTACTGTATGATCACCTGGCATCGCTGGTCATCTGCAGCATT 1229

Db 898 GGGGTGTGCTATGAGACCCCTAGTACTGTATGATCACCTGGCATCGCTGGTCATCTGCAGCATT 957

Qy 1230 TCTTCCCTACTTCATTATTTGGATACACTGCACTTATTTATGCCATCTTATGCTATCTCCTGGTT 1289

Db 958 TCTTCCCTACTTCATTATTTGGATACACTGCACTTATTTATGCCATCTTATGCTATCTCCTGGTT 1017

Qy 1290 TTCCCACTGGCGGTATTTCATGACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAG 1349

Db 1018 TTCCCACTGGCGGTATTTCATGACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAG 1077

Qy 1350 GTCAGCGACCAAGCGCATCCGTGTGACCACTGAAGTTCTCACTTGCAATTAAGCTGATTAATA 1409

Db 1078 GTCAGCGACCAAGCGCATCCGTGTGACCACTGAAGTTCTCACTTGCAATTAAGCTGATTAATA 1137

Qy 1410 ATGTACACATGGGAGAAACCATTTGCAAAAAATCATTTGAAGACCTTAAGAAAGGAAAGG 1469

Db 1138 ATGTACACATGGGAGAAACCATTTGCAGAAATCATTTGAAGACCTTAAGAAAGGAAAGG 1197

Qy 1470 AAACTATTGGAGAAGTGGGGGCTTGTCCAGAGCCTGCACAAAGTATAACCTTGTTCAATCATC 1529

Db 1198 AAACTATTGGAGAAGTGGGGGCTTGTCCAGAGCCTGCACAAAGTATAACCTTGTTCAATCATC 1257

Qy 1530 CCCACAGTGGCCACAGCGGTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACA 1589

Db 1258 CCCACAGTGGCCACAGCGGTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACA 1317

Qy 1590 GCGTCAATGGCCCTTCAGCATGCTGGCCTCCTTGAATCTCCTTCGGCTGTCAGTGTCTTTT 1649

Db 1318 GCGTCAATGGCCCTTCAGCATGCTGGCCTCCTTGAATCTCCTTCGGCTGTCAGTGTCTTTT 1377

Qy 1650 GTGCCTATTGCAGTCAAAAGGTCTCACGAATTTCCAAGTCTGCAGTGTGAGGTTCAAGAAG 1709

Db 1378 GTGCCTATTGCAGTCAAAAGGTCTCACGAATTTCCAAGTCTGCAGTGTGAGGTTCAAGAAG 1437

Qy 1710 TTTTTCCTCCAGGAGAGCCCTGTTTCTATGTCCAGACATTTACAAGACCCCAAGGCT 1769

Db 1438 TTTTTCCTCCAGGAGAGCCCTGTTTCTATGTCCAGACATTTACAAGACCCCAAGGCT 1497

Qy 1770 CTGGTCTTTGAGGAGGCCACCTTGTTCATGGCAACAGACCTGTCCCAGGATCGTCAATGGG 1829

Db 1498 CTGGTCTTTGAGGAGGCCACCTTGTTCATGGCAACAGACCTGTCCCAGGATCGTCAATGGG 1557

Qy 1830 GCACTGGAGCTGGAGAGGAACGGGCATGCTTCTGAGGGGATGACCAAGGCTTAGAGATGCC 1889

Db 1558 GCACTGGAGCTGGAGAGGAACGGGCATGCTTCTGAGGGGATGACCAAGGCTTAGAGATGCC 1617

Qy 1890 CTCGGGCCAGAGGAAGGAAGGAACAGCCCTGGGCCACAGAGTTGCACAAGATCAACCTGGTG 1949

Db 1618 CTCGGGCCAGAGGAAGGAAGGAACAGCCCTGGGCCACAGAGTTGCACAAGATCAACCTGGTG 1677

Qy 1950 GTGTCCAAGGGGATGATGTTAGGGGTCTGCGGCAACACGGGGAGTGGTAAGAGAGCCCTG 2009

Db 1678 GTGTCCAAGGGGATGATGTTAGGGGTCTGCGGCAACACGGGGAGTGGTAAGAGAGCCCTG 1737

Qy 2010 TTGTCAAGCCATCCTGGAGGAGATGCACCTTGTTCGAGGCTCGGTGGGGTGCAGGGAAGC 2069

Db 1738 TTGTCAAGCCATCCTGGAGGAGATGCACCTTGTTCGAGGCTCGGTGGGGTGCAGGGAAGC 1797

QY 2070 CTGGCCTATGTCCCCCAGCAGGCGCTGGATCGTCAAGCGGGAACATCAGGGAGAAACATCCCTC 2129
Db 1798 CTGGCCTATGTCCCCCAGCAGGCGCTGGATCGTCAAGCGGGAACATCAGGGAGAAACATCCCTC 1857
QY 2130 ATGGGAGGCGCATATGACAAGGCCCGGATPACCTCCAGGTGCTCCACTGCTGCCCTGAAT 2189
Db 1858 ATGGGAGGCGCATATGACAAGGCCCGGATPACCTCCAGGTGCTCCACTGCTGCCCTGAAT 1917
QY 2190 CGGGACCTGGAACCTTCTGCCCTTTGGAGACATGACAGAGATTGGAGAGCGGGCGCTCAAC 2249
Db 1918 CGGGACCTGGAACCTTCTGCCCTTTGGAGACATGACAGAGATTGGAGAGCGGGCGCTCAAC 1977
QY 2250 CTCTCTGGGGGCGAGAAACAGAGGATCAGGCTGGCCCGCGCGTCTATTCCGACCGTCAG 2309
Db 1978 CTCTCTGGGGGCGAGAAACAGAGGATCAGGCTGGCCCGCGCGTCTATTCCGACCGTCAG 2037
QY 2310 ATCTACCTGTGGACGACCCCTGTCTGCTGTGGACGCCACCGTGGGGAAGCACATTTTT 2369
Db 2038 ATCTACCTGTGGACGACCCCTGTCTGCTGTGGACGCCACCGTGGGGAAGCACATTTTT 2097
QY 2370 GAGGAGTGCAATTAAGAAAGACACTCAGGGGGAAGACGGTCTGCTCCGTGACCCACGCTG 2429
Db 2098 GAGGAGTGCAATTAAGAAAGACACTCAGGGGGAAGACGGTCTGCTCCGTGACCCACGCTG 2157
QY 2430 CAGTACTTAGAATTTTGTGGCCAGATCATTTTGTGGAAAAATGGAAAAATCTGTGAAAAAT 2489
Db 2158 CAGTACTTAGAATTTTGTGGCCAGATCATTTTGTGGAAAAATGGAAAAATCTGTGAAAAAT 2217
QY 2490 GGAACCTCACAGTGAGTTAATGCAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATG 2549
Db 2218 GGAACCTCACAGTGAGTTAATGCAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATG 2277
QY 2550 CACAAGGAAGCCACTTCGGACATGTTGCAGGACACAGCAAAAGATAGCAGAGAAGCCAAAG 2609
Db 2278 CACAAGGAAGCCACTTCGGACATGCTGCAGGACACAGCAAAAGATAGCAGAGAAGCCAAAG 2337
QY 2610 GTAGAAAGTCAGGCTCTGGCCACCTCCCTGGGAAGAGTCTCTCAACGGAAATGCTGTGCCG 2669
Db 2338 GTAGAAAGTCAGGCTCTGGCCACCTCCCTGGGAAGAGTCTCTCAACGGAAATGCTGTGCCG 2397
QY 2670 GAGCATCAGCTCACACAGGAGGAGGAGATGGAAGAAAGGCTCCTTGAGTTGGAGGGTCTAC 2729
Db 2398 GAGCATCAGCTCACACAGGAGGAGGAGATGGAAGAAAGGCTCCTTGAGTTGGAGGGTCTAC 2457
QY 2730 CACCACCTACATCCAGGCAGCTGGAGGTTACATGGTCTCTTGCAATAATTTCTCTTCGTG 2789
Db 2458 CACCACCTACATCCAGGCAGCTGGAGGTTACATGGTCTCTTGCAATAATTTCTCTTCGTG 2517
QY 2790 GTGCTGATCGTCTCTTAACGATCTTCAGCTTCTGGTGGCTGAGTACTGGTTGGAGCAG 2849
Db 2518 GTGCTGATCGTCTCTTAACGATCTTCAGCTTCTGGTGGCTGAGTACTGGTTGGAGCAG 2577
QY 2850 GGCTCGGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATT 2909
Db 2578 GGCTCGGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATT 2637
QY 2910 GCAGACAATCCTCAACTGTCTCTTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCCTC 2969
Db 2638 GCAGACAATCCTCAACTGTCTCTTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCCTC 2697
QY 2970 ATCTGTGTGGGGTCTGCTCTCTCAGGATTTTACCAAAAGTCAAGGAAGGCATCCACG 3029
Db 2698 ATCTGTGTGGGGTCTGCTCTCTCAGGATTTTACCAAAAGTCAAGGAAGGCATCCACG 2757
QY 3030 GCCCTGCACAACAAGCTCTTCAACAAGGTTTCCGCTGCCCTGAGTTTCTTTGACACC 3089
Db 2758 GCCCTGCACAACAAGCTCTTTAACAAGGTTTCCGCTGCCCTGAGTTTCTTTGACACC 2817
QY 3090 ATCCCAATAGCCGGCTTTTGAACCTGCTCGCAGGGGACTTGGAAACAGCTGGACCAGCTC 3149
Db 2818 ATCCCAATAGCCGGCTTTTGAACCTGCTCGCAGGGGACTTGGAAACAGCTGGACCAGCTC 2877
QY 3150 TTGCCCATCTTTTCAGAGCAGTTTCCCTGGTCTGCTCTTAATGGTGATCGCCGCTCCTGTTG 3209

Db 2878 TTGCCCATCTTTTCAGAGCAGTTTCCCTGGTCTGTCTCTTAATGGTGATCGCCGCTCCTGTTG 2937
QY 3210 ATTGTCAAGTGTGCTGTCTCCATATATCCTGTTAATGGGAGCCATAATCATGTTTATTTGC 3269
Db 2938 ATTGTCAAGTGTGCTGTCTCCATATATCCTGTTAATGGGAGCCATAATCATGTTTATTTGC 2997
QY 3270 TTCAATTTATATATATGATGTTCAAGAAGGCCCATCGGTGTGTTCAAGAGACTGGAGAACTAT 3329
Db 2998 TTCAATTTATATATGATGTTCAAGAAGGCCCATCGGTGTGTTCAAGAGACTGGAGAACTAT 3057
QY 3330 AGCCGGTCTCCTTTATTCTCCCACATCCTCAATTCTCTGCAAGCCCTGAGCTCCATCCAT 3389
Db 3058 AGCCGGTCTCCTTTATTCTCCCACATCCTCAATTCTCTGCAAGCCCTGAGCTCCATCCAT 3117
QY 3390 GTCTATGGAAAAAACTGAAGACTTTCATCAGCCAGTTTAAAGAGGCTGACTGATCGGCAGAAT 3449
Db 3118 GTCTATGGAAAAAACTGAAGACTTTCATCAGCCAGTTTAAAGAGGCTGACTGATCGGCAGAAT 3177
QY 3450 AACTACCTGCTGTTGTTTCTATCTTCCACACAGATGGATGGCATTTGAGGCTGGAGATCATG 3509
Db 3178 AACTACCTGCTGTTGTTTCTATCTTCCACACAGATGGATGGCATTTGAGGCTGGAGATCATG 3237
QY 3510 ACCAACCTTGTGACCTTGGCTGTTGCCCTGTTCTGTTGGCATTTTCCACCCCC 3569
Db 3238 ACCAACCTTGTGACCTTGGCTGTTGCCCTGTTCTGTTGGCATTTTCCACCCCC 3297
QY 3570 TACTCCTTTAAAGTCAATGGCTGTCAACATCGTGTGAGGCTGGCGTCCAGCTTCCAGGCC 3629
Db 3298 TACTCCTTTAAAGTCAATGGCTGTCAACATCGTGTGAGGCTGGCGTCCAGCTTCCAGGCC 3357
QY 3630 ACTGCCCGGATTTGGCTTTGGAGACAGAGGCACAGTTTACACATGGAAGGCACAAGTTGTCCCCAG 3689
Db 3358 ACTGCCCGGATTTGGCTTTGGAGACAGAGGCACAGTTTACACATGGAAGGCACAAGTTGTCCCCAG 3417
QY 3690 TACATGAAGATGTGTCTCGGAAGCTCCTTTACACATGGAAGGCACAAGTTGTCCCCAG 3749
Db 3418 TACATGAAGATGTGTCTCGGAAGCTCCTTTACACATGGAAGGCACAAGTTGTCCCCAG 3477
QY 3750 GGGTGGCCACAGCATGGGAAATCATATTTCAGGATTTACAGATTAACATGAAATACAGAGACAAC 3809
Db 3478 GGGTGGCCACAGCATGGGAAATCATATTTCAGGATTTACAGATTAACATGAAATACAGAGACAAC 3537
QY 3810 ACACCCACCGTGTCTTACGGGCATCAACCTGACCATTCCCGCGCCACGAAGTGGTGGGCATC 3869
Db 3538 ACACCCACCGTGTCTTACGGGCATCAACCTGACCATTCCCGCGCCACGAAGTGGTGGGCATC 3597
QY 3870 GTGGGAAGGACGGGCTCTGGGAAGTCTCTTGGGCATGGCTCTCTTCCGCCCTGGTGGAG 3929
Db 3598 GTGGGAAGGACGGGCTCTGGGAAGTCTCTTGGGCATGGCTCTCTTCCGCCCTGGTGGAG 3657
QY 3930 CCCATGGCAGGCCGGATTCTCATTTACGGCGTGGACATTTGACGATCGGCCCTGGAGGAC 3989
Db 3658 CCCATGGCAGGCCGGATTCTCATTTACGGCGTGGACATTTGACGATCGGCCCTGGAGGAC 3717
QY 3990 TTGCGTCCAAAGCTCTCAGTGATCCCTCAAGATCCAGTGTGCTCTCAGGAACCATCAGA 4049
Db 3718 TTGCGTCCAAAGCTCTCAGTGATCCCTCAAGATCCAGTGTGCTCTCAGGAACCATCAGA 3777
QY 4050 TTCAACCTTAGATCCCTTTTGACCCTGCACACTGACCAGCAGATCTGGGATGCCCTTGAGAGG 4109
Db 3778 TTCAACCTTAGATCCCTTTTGACCCTGCACACTGACCAGCAGATCTGGGATGCCCTTGAGAGG 3837
QY 4110 ACATTCTCTGACCAAGGCCATCTCAAAAGTTCCCCAAAAAGCTGCATACAGATGTGGTGGAA 4169
Db 3838 ACATTCTCTGACCAAGGCCATCTCAAAAGTTCCCCAAAAAGCTGCATACAGATGTGGTGGAA 3897
QY 4170 AACGGTGGAAACTTCTCTGTGGGGGAGAGGCAGCTGCTCTGATTTGCCAGGGCTGTGCTT 4229
Db 3898 AACGGTGGAAACTTCTCTGTGGGGGAGAGGCAGCTGCTCTGATTTGCCAGGGCTGTGCTT 3957
QY 4230 CGAACTCCAAAGATCATCCTTTATCGATGAAGCCACAGCCTCCATTTGACATGGAGACAGAC 4289

Db 3958 CGCAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCTCCATTGACATGGAGACAGAC 4017

Qy 4290 ACCCTGATCCAGCGCACAAATCCGTGAAGCCTTCCAGGGCTGCACCGTGCTCGTCAATTGCC 4349

Db 4018 ACCCTGATCCAGCGCACAAATCCGTGAAGCCTTCCAGGGCTGCACCGTGCTCGTCAATTGCC 4077

Qy 4350 CACCGTGTCACCACCTGTGCTGAACCTGTGACCACATCCTCGTTATGGCAATGGGAAGGTG 4409

Db 4078 CACCGTGTCACCACCTGTGCTGAACCTGTGACCACATCCTCGTTATGGCAATGGGAAGGTG 4137

Qy 4410 GTAGAAATTTGATCGGCCGGAGGTACTGCTGAACTGTGACCACATCCTGGTTCATTGTTGCGAGCCCTC 4469

Db 4138 GTAGAAATTTGATCGGCCGGAGGTACTGCGGAAGAAGCCTGGGTCAATTGTTGCGAGCCCTC 4197

Qy 4470 ATGCCACAGCCACCTTCTTCACTGAGATAAGGAGATGTGGAGACTTTCATGGAGGCTGGCA 4529

Db 4198 ATGCCACAGCCACCTTCTTCACTGAGATAAGGAGATGTGGAGACTTTCATGGAGGCTGGCA 4257

Qy 4530 GCTGAGCTCAGAGGTTTCACACAGGTCAGCTTCGAGGCCACAGTCTGCGACCTTCTTGT 4589

Db 4258 GCTGAGCTCAGAGGTTTCACACAGGTCAGCTTCGAGGCCACAGTCTGCGACCTTCTTGT 4317

Qy 4590 TTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAATGTAGGGGGGTGGGGATTGCTGG 4649

Db 4318 TTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAATGTAGGGGGGTGGGGATTGCTGG 4377

Qy 4650 ATGGAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCGAACCATC 4709

Db 4378 ATGGAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCGAACCATC 4437

Qy 4710 TAAGACATGGGATTTCAGTGATCATGTGGTTCTCCTTTTAACTTACATGCTGAATAATTT 4769

Db 4438 TAAGACATGGGATTTCAGTGATCATGTGGTTCTCCTTTTAACTTACATGCTGAATAATTT 4497

Qy 4770 ATAATAAGGTAAAGCTTATAGTTTCTTGATCTGTGTAGAAAGTGTTCGCAAAATGCTGTAC 4829

Db 4498 ATAATAAGGTAAAGCTTATAGTTTCTTGATCTGTGTAGAAAGTGTTCGCAAAATGCTGTAC 4557

Qy 4830 TGACTTTGTAAAAATATAAACTAAGGAAAACTC 4862

Db 4558 TGACTTTGTAAAAATATAAACTAAGGAAAACTC 4590

RESULT 3

AF352582

LOCUS AF352582 4492 bp mRNA linear PRI 15-MAR-2001

DEFINITION Homo sapiens ATP-binding cassette transporter MRP8 mRNA, complete cds.

ACCESSION AF352582

VERSION AF352582.1 GI:13346480

KEYWORDS .

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 4492)

AUTHORS Bera,T.K., Lee,B.K., Lee,S., Vincent,J. and Pastan,I.H.

TITLE MRP8, a new member of the human ATP-binding cassette (ABC) transporter superfamily which has restricted expression in few tissues

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 4492)

AUTHORS Bera,T.K., Lee,B.K., Lee,S., Vincent,J. and Pastan,I.H.

TITLE Direct Submission

JOURNAL Submitted (23-FEB-2001) Laboratory of Molecular Biology, National Cancer Institute, NIH, 37 Convent Dr., Bethesda, MD 20892, USA

FEATURES

source

1..4492

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="16"

/map="16q12.1"

/tissue_type="breast"

CDS

100..4248

/note="ABC transporter"

/codon_start=1

/product="ATP-binding cassette transporter MRP8"

/protein_id="AAK19755.1"

/db_xref="GI:13346481"

/translation="MTRKRTYWVPNSSGGLVNRGIDIGDDMVSGLIYKITYTLQDGPWS QOERNPEAPGRAVPPWGKYDAALRTMIPFRPKPRPAPQPLDNAGLSYLTVSWLTP LMIQSLRSLRDENTIPPLSVHDASDNVORLHRLWEEVSRRGIEKASVLIWMLRFOR TRLIFDALLGICFIASVLGPILIIIPKILEYSEEQLGNVHVGVLCFALESECVKSL SFSSSWIINQRTAIRFRAAVSSFAFEKLIQFKSVIHITSGEASIFFTGDNVYLFEGVC YGPLVLITCASLVTCSSISYFIIGYTAFAIALCYPLVFPLEVFMTRMAYKAAQHTSEV SDQIRIVTSEVLTCIKLIKMYTWEKPEAKIIEIDLRRKERKLLKCKGLVQSILTSITLFI IPAVATAVWVLIHTSLKLLKTASMAFSMLASLNLRLSVFFVPIAVKGLTNSKSAVMR FKKEFLOESPVFYVQTLDPSKALVFEPEATLSWQQTCPGIVNGALELERNGHASEGMT RPRDALGPEEENSLGPELHKINLVSKGMLGVCGNTSGKSSLSAISLEEMHLLLEG SVGVQGSLAYVPOQAWIVSGNIRENIMCGAYDKARYLOVLHCCSLNRDLELLPFGDM TEIGERGNLSSGQOKRISLARAVYSRQIYILDDPLSAVDHAVGKHIFEICIKKTLR GKTVVQVTHQLYLEFCGQVILLENGKICENGTHSELMQKKGYAQLIQMHKEATSD MLQDTAKIAEKPVESQALATSLSESLNGNAVPEHQLTOEEEMEESGLSNRVYHHYIQ AAGGYMVSCIIFPVVLIJVLTIFFSWWLSYWLQESGTSNSSRESNCGTMADLGNIADN PQLSFYQLVYGLNALLLICVGVCSGGIFTKVKASTALHNKLFNKVFRCPMSFFDTI PIGRLNCFAGDLEQDLQLPIFSEQFLVLSMVIADVLSVLSPIILMGAIIIMVI CFIIYMMFKEAIGVFKRLENYSRSPLFSHILNSLQGLSSIHVYGKTEDEFTISQFKRLTD AONNYLLFLSSRWMALRLEIMTNLVTLAVALLFVAFGISSTPYSFKVMVNIIVLQLA SSFOATARIGLETEAQFTAVERILQYMKMCVSEAPLHMEGTSCPQGWPOHGEILIFQDY HMKYRDNTPTVLHGINLTIRGHEVVGIJVRGTCGSKSSLGMALFRLVEPMAGRILIDGV DICSIGLEDLRKLSVIPQDPVLLSGTIRFNLDPDFDRHTDQOIWDALERTFLTKAISK FPKKLHTDVVENGNFSVGEROLLCIARAVLRNSKIILIDEATASIDMETDTLIQRTI REAFQCGTVLVIAHRVTTVLNCDRILVMGNKVVEFDRPEVLRKKPGSLFAALMATAT SSLR"

BASE COUNT 1073 a 1158 c 1197 g 1064 t

ORIGIN

Query Match 90.3%; Score 4392.8; DB 9; Length 4492;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 4403; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 324 GAAACGAGAATTTGGCAGGAACCTGAAAAATGACTAGGAAGAGGACATACTGGTGCCCAAC 383

Db 73 GAAAAGCAGAAATTGGCAGGAACCTGAAAAATGACTAGGAAGAGGACATACTGGTGCCCAAC 132

Qy 384 TCTTCTGGTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTT 443

Db 133 TCTTCTGGTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTT 192

Qy 444 ATTTATAAAACCTATACTCTCCAAGATGGCCCCCTGGAGTCAGCAAGAGAGAAAATCCTGAG 503

Db 193 ATTTATAAAACCTATACTCTCCAAGATGGCCCCCTGGAGTCAGCAAGAGAGAAAATCCTGAG 252

Qy 504 GCTCCAGGGAGGGCAGCTGTGCCACCCTGGGGGAAGTATGATGCTGCCTTGAGAACCATG 563

Db 253 GCTCCAGGGAGGGCAGCTGTCCCACCCTGGGGGAAGTATGATGCTGCCTTGAGAACCATG 312

Qy 564 ATTCCCTTCCGTCCCAAGCCGAGGTTTCCTGCCCCCAGCCCCCTGGACAATGCTGGCCTG 623

Db 313 ATTCCCTTCCGTCCCAAGCCGAGGTTTCCTGCCCCCAGCCCCCTGGACAATGCTGGCCTG 372

Qy 624 TTCTCCTACCTCACCGTGTGTCATGGCTCACCCCGCTCATGATCCAAAGCTTACGGAGTCGC 683

Db 373 TTCTCCTACCTCACCGTGTGTCATGGCTCACCCCGCTCATGATCCAAAGCTTACGGAGTCGC 432

Qy 684 TTAGATGAGAACACCATCCCTCCACTGTGTCAGTCCATGATGCTTCAGACAAAAATGTCCAA 743

Db 433 TTAGATGAGAACACCATCCCTCCACTGTGTCAGTCCATGATGCTTCAGACAAAAATGTCCAA 492

Qy 744 AGGCTTCACCGCCTTTGGGAAGAAGAAGTCTCAAGGCGAGGGGATTGAAAAAGCTTCAGTG 803

Db 493 AGGCTTCACCGCCTTTGGGAAGAAGAAGTCTCAAGGCGAGGGGATTGAAAAAGCTTCAGTG 552

Qy 804 CTTCTGGTGATGCTGAGGTTCCAGAGAACAAAGGTTGATTTTCGATGCACTTCTGGGCATC 863

Db 553 CTTCTGGTGATGCTGAGGTTCCAGAGAACAAAGGTTGATTTTCGATGCACTTCTGGGCATC 612

QY	864	TGCTTCTGCATTGCCAGTGTACTCGGGCCAATATTGATTATACCAAGATCCTGGAATAT	923
Db	613	TGCTTCTGCATTGCCAGTGTACTCGGGCCAATATTGATTATACCAAGATCCTGGAATAT	672
QY	924	TCAGAAAGAGCAGTTGGGGAATGTTGTCCATGGAGTGGGACTCTGCTTGGCCCTTTTTCCTC	983
Db	673	TCAGAAAGAGCAGTTGGGGAATGTTGTCCATGGAGTGGGACTCTGCTTGGCCCTTTTTCCTC	732
QY	984	TCCGAATGTGGAAGTCTCTGAGTTTCTCCCTTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCT	1043
Db	733	TCCGAATGTGGAAGTCTCTGAGTTTCTCCCTTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCT	792
QY	1044	ATCAGGTTCCGAGCAGCTGTTTTCCTCCTTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCT	1103
Db	793	ATCAGGTTCCGAGCAGCTGTTTTCCTCCTTTGCCCTTTGAGAAGCTCATCCAATTTAAGTCT	852
QY	1104	GTAATACACATCACCTCAGGAGAGGCCATCAGCTTCTTCACCGGTGATGTAACCTACCTG	1163
Db	853	GTAATACACATCACCTCAGGAGAGGCCATCAGCTTCTTCACCGGTGATGTAACCTACCTG	912
QY	1164	TTTGAAGGGTGTGCTATGGACCCCTAGTACTGATCACCTGCGCATCGCTGGTCATCTGC	1223
Db	913	TTTGAAGGGTGTGCTATGGACCCCTAGTACTGATCACCTGCGCATCGCTGGTCATCTGC	972
QY	1224	AGCATTTCTCCTACTTTCATTATTGGATACACTGCATTTATTGCCATCTTATGCTATCTC	1283
Db	973	AGCATTTCTCCTACTTTCATTATTGGATACACTGCATTTATTGCCATCTTATGCTATCCTC	1032
QY	1284	CTGGTTTTCCCACTGGCGGTATTTCATGACAAGAAATGGCTGTGAAGCTCAGCATCACACA	1343
Db	1033	CTGGTTTTCCCACTGGAGGTATTTCATGACAAGAAATGGCTGTGAAGCTCAGCATCACACA	1092
QY	1344	TCTGAGGTCAGCGACCCAGCGCATCCGTGTGACCAAGTGAAGTTCTCACTTGCAATTAAGCTG	1403
Db	1093	TCTGAGGTCAGCGACCCAGCGCATCCGTGTGACCAAGTGAAGTTCTCACTTGCAATTAAGCTG	1152
QY	1404	ATTAAAAATGACACATGGGAGAAACCAATTTGCAAAAAATCATTGAAGACCTTAAGAAAGGAA	1463
Db	1153	ATTAAAAATGACACATGGGAGAAACCAATTTGCAAAAAATCATTGAAGACCTTAAGAAAGGAA	1212
QY	1464	GAAAGGAAACTATTGGAGAAGTGCAGGCTGTCCAGAGCCTGACAAGTATAACCTTGTTTC	1523
Db	1213	GAAAGGAAAGTATTGGAGAAGTGCAGGCTGTCCAGAGCCTGACAAGTATAACCTTGTTTC	1272
QY	1524	ATCATCCCCACAGTGGCCACAGCGGTCTGGGGTTCTCATCCACACATCCTTAAGCTGAAA	1583
Db	1273	ATCATCCCCCAGTGGCCACAGCGGTCTGGGGTTCTCATCCACACATCCTTAAGCTGAAA	1332
QY	1584	CTCACAGCGTCAATGGCCCTTCAGCATGCTGGCCCTCCTTGAATCTCCTTCGGCTGTCAGTG	1643
Db	1333	CTCACAGCGTCAATGGCCCTTCAGCATGCTAGCCCTCCTTGAATCTCCTTCGGCTGTCAGTG	1392
QY	1644	TTCTTTTGGCCTATTGCAGTCAAAGGTCTCACGAATTCGAAGTCTGCAGTGATGAGGTTTC	1703
Db	1393	TTCTTTTGGCCTATTGCAGTCAAAGGTCTCACGAATTCGAAGTCTGCAGTGATGAGGTTTC	1452
QY	1704	AAGAAGTTTTTCTCCAGGAGAGCCCTGTTTTCTATGTCCAGACATTAACAAGACCCCCAGC	1763
Db	1453	AAGAAGTTTTTCTCCAGGAGAGCCCTGTTTTCTATGTCCAGACATTAACAAGACCCCCAGC	1512
QY	1764	AAAGCTCTGGTCTTTTGAGGAGGCCACCTTGTCATGGCAACAGACCTGTCCCGGATCGTC	1823
Db	1513	AAAGCTCTGGTCTTTTGAGGAGGCCACCTTGTCATGGCAACAGACCTGTCCCGGATCGTC	1572
QY	1824	AATGGGGCACTGGAGCTGGAGAGGAACGGGCATGCTTCTGAGGGGATGACCAAGGCCTAGA	1883
Db	1573	AATGGGGCACTGGAGCTGGAGAGGAACGGGCATGCTTCTGAGGGGATGACCAAGGCCTAGA	1632
QY	1884	GATGCCCTCGGGCCAGAGGAAGGAACAGCCCTGGGCCCCAGAGTTGCACAAGATCAAC	1943
Db	1633	GATGCCCTCGGGCCAGAGGAAGGAACAGCCCTGGGCCCCAGAGTTGCACAAGATCAAC	1692

QY	1944	CTGGTGGTGTCCAAGGGGATGATGTTAGGGGTCTCGGGCAACACGGGAGTGGTAAGAGC	2003
Db	1693	CTGGTGGTGTCCAAGGGGATGATGTTAGGGGTCTCGGGCAACACGGGAGTGGTAAGAGC	1752
QY	2004	AGCCTGTTGTGAGCCATCCTCGAGGAGATGCACATTGCTCGAGGGCTCGTGGGGTGCAG	2063
Db	1753	AGCCTGTTGTGAGCCATCCTCGAGGAGATGCACATTGCTCGAGGGCTCGTGGGGTGCAG	1812
QY	2064	GGAAGCCCTGSCCTATGTCCCCCAGCAGGSCCTGGATCGTCAGCGGGACATCAGGGAGAAC	2123
Db	1813	GGAAGCCCTGSCCTATGTCCCCCAGCAGGSCCTGGATCGTCAGCGGGACATCAGGGAGAAC	1872
QY	2124	ATCCTCATGGGAGGCGCATATGACAAGGCCCGCATACCTCCAGGTGCTCCACTGCTCTCC	2183
Db	1873	ATCCTCATGGGAGGCGCATATGACAAGGCCCGCATACCTCCAGGTGCTCCACTGCTCTCC	1932
QY	2184	CTGAATCGGACCTTGAACCTTCTGCCCTTTGGAGACATGACAGAGATTGGAGAGCGGGC	2243
Db	1933	CTGAATCGGACCTTGAACCTTCTGCCCTTTGGAGACATGACAGAGATTGGAGAGCGGGC	1992
QY	2244	CTCAACCTCTCTGGGGGCGAGAAACAGAGGATCAGCCTGGCCCGCGCTCTATTTCCGAC	2303
Db	1993	CCCAACCTCTCTGGGGGCGAGAAACAGAGGATCAGCCTGGCCCGCGCTCTATTTCCGAC	2052
QY	2304	CGTCAGATCTACCTGCTGGACGACCCCTGTCTGCTGTGGACGCCACCTGGGGAAGCAC	2363
Db	2053	CGTCAGATCTACCTGCTGGACGACCCCTGTCTGCTGTGGACGCCACCTGGGGAAGCAC	2112
QY	2364	ATTTTGTAGGAGTGCAATTAAGAAGACACTCAGGGGGAAGACGGTCGTCCTGGTGACCCAC	2423
Db	2113	ATTTTGTAGGAGTGCAATTAAGAAGACACTCAGGGGGAAGACGGTCGTCACAGGTGACCCAC	2172
QY	2424	CAGCTGCAGTACTTAGAATTTTGTGGCCAGATCATTTTGTGGAAATGGGAAAAATCTGT	2483
Db	2173	CAGCTGCAGTACTTAGAATTTTGTGGCCAGGTCAATTTTGTGGAAATGGGAAAAATCTGT	2232
QY	2484	GAAAATGGAACTCACAGTGAGTTAATGCGAGAAAAAGGGGAAATATGCCCAACTTATCCAG	2543
Db	2233	GAAAATGGAACTCACAGTGAGTTAATGCGAGAAAAAGGGGAAATATGCCCAACTTATCCAG	2292
QY	2544	AAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGGACACAGCAAGATAGCAGAGAAG	2603
Db	2293	AAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGGACACAGCAAGATAGCAGAGAAG	2352
QY	2604	CCAAAGGTAGAAAAGTCAGGCTCTGGCCACCTCCCTGGAAAGAGTCTCTCAACGGAAAATGCT	2663
Db	2353	CCAAAGGTAGAAAAGTCAGGCTCTGGCCACCTCCCTGGAAAGAGTCTCTCAACGGAAAATGCT	2412
QY	2664	GTCCCGGAGCATCAGCTCACACAGGAGGAGGATGGAAGAGGCTCCTTGAGTTGGAGG	2723
Db	2413	GTCCCGGAGCATCAGCTCACACAGGAGGAGGATGGAAGAGGCTCCTTGAGTTGGAGG	2472
QY	2724	GTCTACCAACCACTACATCCAGGCAAGCTGGAGGTTACATGGTCTCTTGCATAAATTTCTTC	2783
Db	2473	GTCTACCAACCACTACATCCAGGCAAGCTGGAGGTTACATGGTCTCTTGCATAAATTTCTTC	2532
QY	2784	TTCTGTTGCTGATCGTCTTCTTTAACCATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTG	2843
Db	2533	TTTGTGGTGTGATCGTCTTCTTTAACCATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTG	2592
QY	2844	GAGCAGGCTCGGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATGGCAGACCTGGGC	2903
Db	2593	GAGCAGGCTCGGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATGGCAGACCTGGGC	2652
QY	2904	AACATTGCAGACAAATCCTCAACTGTCTTCTACCAGCTGGTGTACGGGCTCAACGCCCTG	2963
Db	2653	AACATTGCAGACAAATCCTCAACTGTCTTCTACCAGCTGGTGTACGGGCTCAACGCCCTG	2712
QY	2964	CTCCTCATCTGTGTGGGGGTCTGCTCCTCAGGGATTTTCACCAAAAGTCACGAGGAAGCA	3023
Db	2713	CTCCTCATCTGTGTGGGGGTCTGCTCCTCAGGGATTTTCACCAAAAGTCACGAGGAAGCA	2772
QY	3024	TCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTTTTCCGCTGCCCCATGAGTTTCTTT	3083

Db 2773 TCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTTCCGGCTGCCCATGAGTTCTTT 2832

QY 3084 GACACCATCCCAATAGGCCGGCTTTTGAACTGCTTCGCAGGGGACTTGGAACAGCTGGAC 3143

Db 2833 GACACCATCCCAATAGGCCGGCTTTTGAACTGCTTCGCAGGGGACTTGGAACAGCTGGAC 2892

QY 3144 CAGCTCTTGCCCATCTTTTCAGAGCAGTTTCCTGGTCTCTCCTTAATGGTGATCGCCGTC 3203

Db 2893 CAGCTCTTGCCCATCTTTTCAGAGCAGTTTCCTGGTCTCTCCTTAATGGTGATCGCCGTC 2952

QY 3204 CTGTTGATTGTCAGTGCTGTCTCCATATATCCTGTTAATGGGAGCCCATAAATCATGGTT 3263

Db 2953 CTGTTGATTGTCAGTGCTGTCTCCATATATCCTGTTAATGGGAGCCCATAAATCATGGTT 3012

QY 3264 ATTTGCTTCAATTTATTATATGATGTTCAAGAAGGCCATCGGTGTGTTCAAGAGACTGGAG 3323

Db 3013 ATTTGCTTCAATTTATTATATGATGTTCAAGGAGGCCATCGGTGTGTTCAAGAGACTGGAG 3072

QY 3324 AACTATAGCCGGTCTCCTTTATTTCTCCACATCCTCAATTTCTTGCAAGGCCTGAGCTCC 3383

Db 3073 AACTATAGCCGGTCTCCTTTATTTCTCCACATCCTCAATTTCTTGCAAGGCCTGAGCTCC 3132

QY 3384 ATCCATGTCTATGGAAAACTGAAGACTTTCATCAGCCAGTTTAAAGAGGCTGACTGATCG 3443

Db 3133 ATCCATGTCTATGGAAAACTGAAGACTTTCATCAGCCAGTTTAAAGAGGCTGACTGATCG 3192

QY 3444 CAGAATAACTACCTGCTGTTGTTTCTATCTTCCACACGATGGATGGCATTGAGGCTGGAG 3503

Db 3193 CAGAATAACTACCTGCTGTTGTTTCTATCTTCCACACGATGGATGGCATTGAGGCTGGAG 3252

QY 3504 ATCATGACCAACCTTGTCACCTTGACCTTGTTGCCCTGTTCGTGGCTTTTGGCATTTCCCTCC 3563

Db 3253 ATCATGACCAACCTTGTCACCTTGACCTTGCCCTGTTCGTGGCTTTTGGCATTTCCCTCC 3312

QY 3564 ACCCCCTACTCCTTTAAAGTCATGGCTGTCAACATFCGTGCTGCAGCTGGCGTCCAGCTTC 3623

Db 3313 ACCCCCTACTCCTTTAAAGTCATGGCTGTCAACATCGTGTGCAGCTGGCGTCCAGCTTC 3372

QY 3624 CAGGCCACTGCCCGGATTGGCTTGAGACAGAGGCACAGTTTACGGCTGTAGAGAGGATA 3683

Db 3373 CAGGCCACTGCCCGGATTGGCTTGAGACAGAGGCACAGTTTACGGCTGTAGAGAGGATA 3432

QY 3684 CTGCAGTACATGAAGATGTGTCTCGGAAGCTCCTTTACACATGGAAGGCACAAAGTTGT 3743

Db 3433 CTGCAGTACATGAAGATGTGTCTCGGAAGCTCCTTTACACATGGAAGGCACAAAGTTGT 3492

QY 3744 CCCCAGGGGTGGCCACAGCATGGGGAATCATATTTACAGGATTATCACATGAAATACAGA 3803

Db 3493 CCCCAGGGGTGGCCACAGCATGGGGAATCATATTTACAGGATTATCACATGAAATACAGA 3552

QY 3804 GACAACACACCCACCGTCTCACGGCATCAACTGACCATCCGCGGCCACGAAGTGGTG 3863

Db 3553 GACAACACACCCACCGTCTCACGGCATCAACTGACCATCCGCGGCCACGAAGTGGTG 3612

QY 3864 GGCATCGTGGGAAGGACGGGCTCTGGGAAGTCTCCTTGGGCATGGCTCTCTCCGCCCTG 3923

Db 3613 GGCATCGTGGGAAGGACGGGCTCTGGGAAGTCTCCTTGGGCATGGCTCTCTCCGCCCTG 3672

QY 3924 GTGGAGCCCATGGCAGGCGGATTTCTCATTTGACGGCGTGGACATTTGCAGCATCGGCCCTG 3983

Db 3673 GTGGAGCCCATGGCAGGCGGATTTCTCATTTGACGGCGTGGACATTTGCAGCATCGGCCCTG 3732

QY 3984 GAGGACTTGCGGTCCAAAGCTCTCAGTGATCCCTCAAGATCCAGTGTCTCTCAGGAACC 4043

Db 3733 GAGGACTTGCGGTCCAAAGCTCTCAGTGATCCCTCAAGATCCAGTGTCTCTCAGGAACC 3792

QY 4044 ATCAGATTTCAACCTAGATCCCTTTGACCGTCACTGACCCAGCAGATCTGGATGCCCTTG 4103

Db 3793 ATCAGATTTCAACCTAGATCCCTTTGACCGTCAACCCAGCAGCAGATCTGGATGCCCTTG 3852

QY 4104 GAGAGGACATTCCTGACCAAGGCCATCTCAAAGTTCCCCAAAAAGCTGCATACAGATGTG 4163

Db 3853 GAGAGGACATTCCTGACCAAGGCCATCTCAAAGTTTCCCCAAAAAGCTGCATACAGATGTG 3912

QY 4164 GTGGAAAAACGGTGGAAACTTCTGTGGGGGAGAGCAGCTGCTCTGCATTTGCCAGGGCT 4223

Db 3913 GTGGAAAAACGGTGGAAACTTCTGTGGGGGAGAGCAGCTGCTCTGCATTTGCCAGGGCT 3972

QY 4224 GTGCTTCGCAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCTCCATTTGACATGGAG 4283

Db 3973 GTGCTTCGCAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCTCCATTTGACATGGAG 4032

QY 4284 ACAGACACCCTGATCCAGCCGCAAAATCCCGTGAAGCCTTCCAGGGCTGCACCGTCTCGTC 4343

Db 4033 ACAGACACCCTGATCCAGCGCAAAATCCCGTGAAGCCTTCCAGGGCTGCACCGTCTCGTC 4092

QY 4344 ATTGCCCCACCGTGTACCACTGTGCTGAACCTGTGACCACATCCTGGTTATGGGCAATGGG 4403

Db 4093 ATTGCCCCACCGTGTACCACTGTGCTGAACCTGTGACCGCATCCTGGTTATGGGCAATGGG 4152

QY 4404 AAGTGGTAGAAATTTGATCGGCCGAGGTACTGCGGAAGAACCTTGGGTCAITGTTTCGCA 4463

Db 4153 AAGTGGTAGAAATTTGATCGGCCGAGGTACTGCGGAAGAACCTTGGGTCAITGTTTCGCA 4212

QY 4464 GCCCTCATGGCCACAGCACTTCTTCACTGAGATAAGGAGATGTGGAGACTTTCATTCGAGG 4523

Db 4213 GCCCTCATGGCCACAGCACTTCTTCACTGAGATAAGGAGATGTGGAGACTTTCATTCGAGG 4272

QY 4524 CTGGCAGCTGAGCTCAGAGGTTTCACACAGGTGCAGCTTCGAGGCCACAGTCTCGGACCT 4583

Db 4273 CTGGCAGCTGAGCTCAGAGGTTTCACACAGCTGCAGCTTCGAGGCCACAGTCTCGGACCT 4332

QY 4584 TCTTGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAATGTAGGGGGGTGGGAT 4643

Db 4333 TCTTGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAATGTAGGGGGGTGGGAT 4392

QY 4644 TGCTGGATGGAACCCCTGGAATAGGTAAGTGGCTCTCAAGACCTTAGAACCCACAGA 4703

Db 4393 TGCTGGATGGAACCCCTGGAATAGGTAAGTGGCTCTCAAGACCTTAGAACCCACAGA 4452

QY 4704 ACCATCTAAGACATGGGATTTCAGTGATCATGIGGTTCTCC 4743

Db 4453 ACCATCTAAGACATGGGATTTCAGTGATCATGIGGTTCTCC 4492

RESULT 4
AX468400
LOCUS AX468400 4638 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 1 from Patent WO216589.
ACCESSION AX468400
VERSION AX468400.1 GI:21901253
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Curtis,R.A.
TITLE Human ABC transporter family member and uses thereof
JOURNAL Patent: WO 0216589-A 1 28-FEB-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source location/Qualifiers
1. .4638
/organism="Homo sapiens"
/db_xref="taxon:9606"
201..4283
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD42559.1"
/db_xref="GI:21901254"
/translation="MTRKRTYWPNSSGGLVNRGIDIGDDMVSGLIYKTYTLQDGPWS
OOERNPEAPGRAAPPWKYDAALRTMIPFRPKRFPAPQPLDNAGLFSYLTVSWLTP
LMIQSLRSLDENTIPPLSVHDASDKNVORHLRWEDEVSRRIEKASVLLVMLRFQR
TRLIFDALLGICFCIASVLGPILIIIPKILEYSEEQGNVHVHGLCFALFLSECVKSL
SFSSSWIINORTAIRRAAVSSFAFEKLIQFKSVIHITSGEGGDIQAHQLAVLQAI

FTGDVNYLPEGVVCYGPLVLITCASLVICSISSYFIIGYTAFAIAILCYLLVFPPLAVFMT
RMAVKAQHHTSEVSDQIRIRVTSEVLTCIKLIKMYTWKPEPAKIIIEGMESLTFCSKPGD
GMAFSMLASLNLLRLSVFFPIAVKGLTNSKSAVMREKFPFLQESPVFYVQTLDPSK
ALVEEATLSWQOTCPGIVNGALELRNGHASEGMTRPDALGPPEEGNSLGPELHKI
NLVVGKMMGLGVCNGTSGKSSLSAILEEMHLEGSVGQSLAYVPPQAWIVSGNI
RENILMGAYDKARYLOVLHCCSLNRDLELLPFGDMTEIGERGLNLSGGQKQRIULAR
AVYSDROIYLLDDPLSAVDHVGKHIFEICKIKTLRGTKVLVTHQLQYLEFCGOIIL
LENGKICENGTHSELMOQKGYAQLIOKMHEKATSDMLQDTAKIAEKPKVESQALATS
LEESLGNNAVPEHOLTQEEEMEESLSWRVHHYIQAAAGYMVSCIIFFVVLIVFLT
IFSFWLSYWLEQSGTNSRESNGTMADLGNIAQNPQSLFQYQLVYGLNALLLICVGV
CSSGIFTKVTRKASTALHNKLPFNKVFRCPMSEFFDTIPIGRLNLCFAGDLEQLDQLLPI
FSEQFLVLSLMVIAVLLIVSVLSPIILLMGAIIMVICFIYMMFKKAIGVFKRLNYS
RSPFLSHILNSLOGLSSIHVYKTEDEFISQFKRLTDAQNNYLLFLSSTRWMALREI
MTNLVTLAVALFVAFGISSTPYSEFKVMVAVNIVLQASSFOATARIGLETEAQFTAVER
ILQYMKVCVSEAPLHMEGTSCPQGWPOHGEIIFQDYHMKYRDNTPTVLHGINTLRGH
EVVGIVGRTSGKSSGLMALFRLVEPDMAGRILIDGVDICSIGLEDLRSKLSVIPQDPV
LLSGTIRFNLDPFDRHTDQIWDALERTFLTKAISKFPKKLHTDVVENGGNFSVGERQ
LLCIRARAVLRNSKIILIDEATASIDMETDTLIQRTIREAFQGCITVLVIAHRVTTVLNC
DHILVMGNGKVVEFDRPEVLRKKPGSLFAALMATATSSLR"

BASE COUNT 1106 a 1172 c 1223 g 1137 t
ORIGIN

Query Match 90.3%; Score 4391.4; DB 6; Length 4638;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 4578; Conservative 0; Mismatches 21; Indels 144; Gaps 3;

QY	151	GATGTTTAAAAAGAGGGATCAAGCACAGGCTAAGGAGAGGAAAGAGCAGGCACCCAAACC	210
Db	1	GATGTTTAAAAAGAGGGATCAAGCACAGGCTAAGGAGAGGAAAGAGCAGGCACCCAAACC	60
QY	211	TCTGCATGGCCCCCAATATGCTCCCTGCAGGGTAGTGCCCCCTCTTCTGGCTGCTCAAGGC	270
Db	61	TCTGCATGGCCCCCAATATGCTCCCTGCAGGGTAGTGCCCCCTCTTCTGGCTGCTCAAGGC	120
QY	271	GAGATCTAAGCTTCTTCTAACTCCTGCTGTCTTTTTCATATTCTCTGATTCTGGGAAACGA	330
Db	121	GAGATCTAAGCTTCTTCTAACTCCTGCTGTCTTTTTCATATTCTCTGATTCTGGGAAACGA	180
QY	331	AGAATTGGCAGGAACCTGAAAATGACTAGGAAGAGGACATACTGGTGCCCCAACTCTTCTG	390
Db	181	AGAATTGGCAGGAACCTGAAAATGACTAGGAAGAGGACATACTGGTGCCCCAACTCTTCTG	240
QY	391	GTGGCCTCGTGAATCGTGGCATCGACATPAGGCGATGACATGGTTTCAGGACTTATTTATA	450
Db	241	GTGGCCTCGTGAATCGTGGCATCGACATPAGGCGATGACATGGTTTCAGGACTTATTTATA	300
QY	451	AAACCTATACTCTCCAAGATGGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAG	510
Db	301	AAACCTATACTCTCCAAGATGGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAG	360
QY	511	GGAGGGCAGCTGTCCCACCGTGGGGGAAGTATGATGCTGCCCTTGAGAACCATGATTCCT	570
Db	361	GGAGGGCAGCTGTCCCACCGTGGGGGAAGTATGATGCTGCCCTTGAGAACCATGATTCCT	420
QY	571	TCCGTCCCAAGCCGAGGTTTCTCTGCCCCCCAGCCCCCTGGACAATGCTGGCCCTGTTCTCCT	630
Db	421	TCCGTCCCAAGCCGAGGTTTCTCTGCCCCCCAGCCCCCTGGACAATGCTGGCCCTGTTCTCCT	480
QY	631	ACCTCACCGTGTATGGCTCACCCCGCTCATGATCCTCAAGGCTTACGGAGTCGCTTAGATG	690
Db	481	ACCTCACCGTGTATGGCTCACCCCGCTCATGATCCTCAAGGCTTACGGAGTCGCTTAGATG	540
QY	691	AGAACACCATCCCTCCACTGTCACTGCCATGATGCCCTCAGACAAAAATGTCCAAAGGCTTC	750
Db	541	AGAACACCATCCCTCCACTGTCACTGCCATGATGCCCTCAGACAAAAATGTCCAAAGGCTTC	600
QY	751	ACCGCCTTTGGGAAGAAGAAGTCTCAAGCGGAGGATTTGAAAAAGCTTCAGTGCTTCTGG	810
Db	601	ACCGCCTTTGGGAAGAAGAAGTCTCAAGCGGAGGATTTGAAAAAGCTTCAGTGCTTCTGG	660
QY	811	TGATGCTGAGGTTCCAGAGAAACAAGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTCT	870
Db	661	TGATGCTGAGGTTCCAGAGAAACAAGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTCT	720

QY	871	GCATTGCCAGTGTACTCGGGCCCAATATTGATTATACCAAAGATCTCTGGAATATTCAGAAG	930
Db	721	GCATTGCCAGTGTACTCGGGCCCAATATTGATTATACCAAAGATCTCTGGAATATTCAGAAG	780
QY	931	AGCAGTTGGGGAATGTTGTCCATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAAT	990
Db	781	AGCAGTTGGGGAATGTTGTCCATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAAT	840
QY	991	GTGTGAAGTCTCTGAGTTTCTCCTCCAGTTGGATCATCAACCACGACAGCCATCAGGT	1050
Db	841	GTGTGAAGTCTCTGAGTTTCTCCTCCAGTTGGATCATCAACCACGACAGCCATCAGGT	900
QY	1051	TCCGAGCAGCTGTTTCCCTCCCTTGGCTTTGAGAAGCTCATCCAAATTAAGTCTGTAATAC	1110
Db	901	TCCGAGCAGCTGTTTCCCTCCCTTGGCTTTGAGAAGCTCATCCAAATTAAGTCTGTAATAC	960
QY	1111	ACATCACCTCAGGAG-----AGGCCA	1131
Db	961	ACATCACCTCAGGAGAGGGAGGTGACATCTGTGCCCATCAACTTGCTGTCTTGCAGGCCA	1020
QY	1132	TCAGCTTCTTACCCTGATGTAACTACCTGTTTGAAGGGGTGCTGCTATGGACCCCTAG	1191
Db	1021	TCAGCTTCTTACCCTGATGTAACTACCTGTTTGAAGGGGTGCTGCTATGGACCCCTAG	1080
QY	1192	TACTGATCACCTGCGCATCGCTGGTCATCTGCAGCATTTCTTCCTACTTCATTATTGGAT	1251
Db	1081	TACTGATCACCTGCGCATCGCTGGTCATCTGCAGCATTTCTTCCTACTTCATTATTGGAT	1140
QY	1252	ACACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCACCTGGCGGTATTCTATGA	1311
Db	1141	ACACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCACCTGGCGGTATTCTATGA	1200
QY	1312	CAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGAGGACCAGCGCATCCGTG	1371
Db	1201	CAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTGAGGACCAGCGCATCCGTG	1260
QY	1372	TGACCAGTGAAGTTCTCACTTGCAATTAAGCTGATTAAAAATGTACACATGGGAGAAACCAT	1431
Db	1261	TGACCAGTGAAGTTCTCACTTGCAATTAAGCTGATTAAAAATGTACACATGGGAGAAACCAT	1320
QY	1432	TTGCAAAAATCATTTGAAGACCTTAAGAAGGAAGAAAGAACTATTGGAGAAAGTGGGGC	1491
Db	1321	TTGCAAAAATCATTTGAAG-----	1338
QY	1492	TTGTCCAGAGCCTGCACAAAGTATAACCTTGTTTCATCATCCCCACAGTGGCCACAGCGGTCT	1551
Db	1339	-----GTATGAAAAGTCTGACTTCTGCTCCAAACCTGGTGATGGC-----	1379
QY	1552	GGGTTCTCATCCACACATCCTTAAAGCTGAAAACCTCACAGCGTCAATGGCCTTCAGCATGC	1611
Db	1380	-----ATGGCCTTCAGCATGC	1395
QY	1612	TGGCCTCCTTGAATCTCCTTCGGCTGTCAGTGTTCCTTGTGCCTATTGCAGTCAAAGGTC	1671
Db	1396	TGGCCTCCTTGAATCTCCTTCGGCTGTCAGTGTTCCTTGTGCCTATTGCAGTCAAAGGTC	1455
QY	1672	TCACGAATTCCAAAGTCTGCAGTGTGAGGTTCAAGAAAGTTTTCCTCCAGGAGAGCCCTG	1731
Db	1456	TCACGAATTCCAAAGTCTGCAGTGTGAGGTTCAAGAAAGTTTTCCTCCAGGAGAGCCCTG	1515
QY	1732	TTTTCTATGTCCAGACATTACAAGACCCCGGATCGTCAATGGGGCTTGTGCTTTGAGGAGGCCACCT	1791
Db	1516	TTTTCTATGTCCAGACATTACAAGACCCCGGATCGTCAATGGGGCTTGTGCTTTGAGGAGGCCACCT	1575
QY	1792	TGTCATGGCAACAGACCTGTCCCGGATCGTCAATGGGGCTTGTGAGAGTGGAGAGGAAACG	1851
Db	1576	TGTCATGGCAACAGACCTGTCCCGGATCGTCAATGGGGCTTGTGAGAGTGGAGAGGAAACG	1635
QY	1852	GGCATGCTTCTGAGGGGATGACCAAGGCTTAGAGATGCCCTCGGSCCAGAGGAAGGGA	1911
Db	1636	GGCATGCTTCTGAGGGGATGACCAAGGCTTAGAGATGCCCTCGGSCCAGAGGAAGGGA	1695

Qy	1912	ACAGCCTGGGCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCTCAAGGGGATGATGTTAG	1971
Db	1696	ACAGCCTGGGCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCTCAAGGGGATGATGTTAG	1755
Qy	1972	GGTCTCGGGCAACACCGGAGTGGTAAGAGCAGCCCTGTGTGTCAGCCATCCTGGAGGAGA	2031
Db	1756	GGTCTCGGGCAACACCGGAGTGGTAAGAGCAGCCCTGTGTGTCAGCCATCCTGGAGGAGA	1815
Qy	2032	TGCACCTGCTCGAGGGCTCGTGGGGGTGCAGGGAAGCCTGGCCATGTCCCCAGCAGG	2091
Db	1816	TGCACCTGCTCGAGGGCTCGTGGGGGTGCAGGGAAGCCTGGCCATGTCCCCAGCAGG	1875
Qy	2092	CCTGGATCGTCAGCGGGAACATCAGGGAGAACATCCTCATGGGAGGCGCATATGACAAGG	2151
Db	1876	CCTGGATCGTCAGCGGGAACATCAGGGAGAACATCCTCATGGGAGGCGCATATGACAAGG	1935
Qy	2152	CCCGATACCTCCAGGTGCTCCACTGCTGCTGCCCTGAATCGGGACCTGGAACTTCTGCCCT	2211
Db	1936	CCCGATACCTCCAGGTGCTCCACTGCTGCTGCCCTGAATCGGGACCTGGAACTTCTGCCCT	1995
Qy	2212	TTGGAGACATGACAGAGATTTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGGCAGAAACAGA	2271
Db	1996	TTGGAGACATGACAGAGATTTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGGCAGAAACAGA	2055
Qy	2272	GGATCAGCCTGGCCCCGGCCGTCTATTCCGACCCTCAGATCTACCTGCTGGACGCCCCCC	2331
Db	2056	GGATCAGCCTGGCCCCGGCCGTCTATTCCGACCCTCAGATCTACCTGCTGGACGCCCCCC	2115
Qy	2332	TGTCTGCTGTGGACGCCACGCTGGGGAAGCACATTTTGTAGGAGTGCATTAAGAAGACAC	2391
Db	2116	TGTCTGCTGTGGACGCCACGCTGGGGAAGCACATTTTGTAGGAGTGCATTAAGAAGACAC	2175
Qy	2392	TCAGGGGGAAGACGGTCGTCCTCGGTGACCCACCAGCTGCAGTACTTAGAATTTGTGGCC	2451
Db	2176	TCAGGGGGAAGACGGTCGTCCTCGGTGACCCACCAGCTGCAGTACTTAGAATTTGTGGCC	2235
Qy	2452	AGATCATTTTGTGGAAATGGGAAAATCTGTGAAATGGAACCTCACAGTGAGTTAATGC	2511
Db	2236	AGATCATTTTGTGGAAATGGGAAAATCTGTGAAATGGAACCTCACAGTGAGTTAATGC	2295
Qy	2512	AGAAAAGGGGAAATATGCCCAACTTATCCAGAAAGATGCACAAGGAAGCCACTTCGGACA	2571
Db	2296	AGAAAAGGGGAAATATGCCCAACTTATCCAGAAAGATGCACAAGGAAGCCACTTCGGACA	2355
Qy	2572	TGTTGCAGGACACACGCAAAAGATAGCAGAGAAGCCAAAGGTAGAAAGTCAGGCTCTGGCCA	2631
Db	2356	TGTTGCAGGACACACGCAAAAGATAGCAGAGAAGCCAAAGGTAGAAAGTCAGGCTCTGGCCA	2415
Qy	2632	CCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCGGAGCATCAGCTCACACAGGAGG	2691
Db	2416	CCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCGGAGCATCAGCTCACACAGGAGG	2475
Qy	2692	AGGAGATGGAAGAAGGCTCCTTGAGTTGGAGGGTCTACCACCACCTACATCCAGGCAGCTG	2751
Db	2476	AGGAGATGGAAGAAGGCTCCTTGAGTTGGAGGGTCTACCACCACCTACATCCAGGCAGCTG	2535
Qy	2752	GAGGTTACATGGTCTCTGCATAAATTTCTCTCTGTTGGTGCTGATCGTCTCTTAACGA	2811
Db	2536	GAGGTTACATGGTCTCTGCATAAATTTCTCTCTGTTGGTGCTGATCGTCTCTTAACGA	2595
Qy	2812	TCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGCAGCC	2871
Db	2596	TCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGCAGCC	2655
Qy	2872	GAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCAGACAAATCCTCAACTGTCCCT	2931
Db	2656	GAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTCAGACAAATCCTCAACTGTCCCT	2715
Qy	2932	TCTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGGTCTGCTCCT	2991
Db	2716	TCTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGGTCTGCTCCT	2775
Qy	2992	CAGGGATTTTCACCAAAAGTCACGAGGAAGGCATCCACGGCCCTGCACAAACAGCTCTTCA	3051

Db	2776	CAGGGATTTTCACCAAAAGTCACGAGGAAGGCATCCACGGCCCTGCACAACAAGCTCTTCA	2835
Qy	3052	ACAAGGTTTTCCGCTGCCCCATGAGTTTTCTTTTGACACCATCCCAATAGGCCGGCTTTGA	3111
Db	2836	ACAAGGTTTTCCGCTGCCCCATGAGTTTTCTTTTGACACCATCCCAATAGGCCGGCTTTGA	2895
Qy	3112	ACTGCTTCGCAGGGGACTTGGAAACAGCTGGACCAGCTCTTTGCCCATCTTTTCAGAGCAGT	3171
Db	2896	ACTGCTTCGCAGGGGACTTGGAAACAGCTGGACCAGCTCTTTGCCCATCTTTTCAGAGCAGT	2955
Qy	3172	TCCTGGTCCCTGTCCTTAATGGTGATCGCCGTCCTGTTCATGTTCAGTGTGCTGTCCAT	3231
Db	2956	TCCTGGTCCCTGTCCTTAATGGTGATCGCCGTCCTGTTCATGTTCAGTGTGCTGTCCAT	3015
Qy	3232	ATATCCTGTTAATGGGAGCCATAAATCATGGTTAATTTGCTTCATTTATTATAJGATGTTCA	3291
Db	3016	ATATCCTGTTAATGGGAGCCATAAATCATGGTTAATTTGCTTCATTTATTATAJGATGTTCA	3075
Qy	3292	AGAAGGCCAJTCGGTGTGTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATCTCCC	3351
Db	3076	AGAAGGCCAJTCGGTGTGTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTTATCTCCC	3135
Qy	3352	ACATCCTCAATTCTCTGCAAGCCCTGAGCTCCATCCATGCTCTATGGAAAAAACTGAAGACT	3411
Db	3136	ACATCCTCAATTCTCTGCAAGCCCTGAGCTCCATCCATGCTCTATGGAAAAAACTGAAGACT	3195
Qy	3412	TCATCAGCCAGTTTAAGAGGCTGACTGATGGCAGAAATAACTACCTGCTGTTGTTCTAT	3471
Db	3196	TCATCAGCCAGTTTAAGAGGCTGACTGATGGCAGAAATAACTACCTGCTGTTGTTCTAT	3255
Qy	3472	CTTCCACACGATGGATGGATTGAGGCTGGAGATCATGACCAACCTTTGTGACCTTGGCTG	3531
Db	3256	CTTCCACACGATGGATGGATTGAGGCTGGAGATCATGACCAACCTTTGTGACCTTGGCTG	3315
Qy	3532	TTGCCCTGTTTCGTGGCTTTTGGCATTTCTCCACCCCTACTCCTTTAAAGTCATGGCTG	3591
Db	3316	TTGCCCTGTTTCGTGGCTTTTGGCATTTCTCCACCCCTACTCCTTTAAAGTCATGGCTG	3375
Qy	3592	TCAACATCGTGTGCAGCTGCGCTCCAGCTTCCAGGCCACTGCCCGGATTTGGCTTGAGA	3651
Db	3376	TCAACATCGTGTGCAGCTGCGCTCCAGCTTCCAGGCCACTGCCCGGATTTGGCTTGAGA	3435
Qy	3652	CAGAGGCACAGTTCACGGCTCTAGAGAGGATACTGCAGTACATGAAGATGTGTGTCCTGG	3711
Db	3436	CAGAGGCACAGTTCACGGCTCTAGAGAGGATACTGCAGTACATGAAGATGTGTGTCCTGG	3495
Qy	3712	AAGCTCCTTTACACATGGAAGGCACAAGTTGTCCCCAGGGGTGGCCACAGCATGGGGAAA	3771
Db	3496	AAGCTCCTTTACACATGGAAGGCACAAGTTGTCCCCAGGGGTGGCCACAGCATGGGGAAA	3555
Qy	3772	TCATATTTTCAGGATTTACATGAAATACAGAGACAACACCCACCTGCTGCTTCACGGCA	3831
Db	3556	TCATATTTTCAGGATTTACATGAAATACAGAGACAACACACCCACCTGCTGCTTCACGGCA	3615
Qy	3832	TCAACCTGACCATCCGGGCCACGAAAGTGGTGGGCACTCGTGGGAAGGACGGGCTCTGGGA	3891
Db	3616	TCAACCTGACCATCCGGGCCACGAAAGTGGTGGGCACTCGTGGGAAGGACGGGCTCTGGGA	3675
Qy	3892	AGTCCTCCTTGGGCATGGCTCTCTTCCGCCCTGGTGGAGCCCCATGGCAGGCCGGATTTCTCA	3951
Db	3676	AGTCCTCCTTGGGCATGGCTCTCTTCCGCCCTGGTGGAGCCCCATGGCAGGCCGGATTTCTCA	3735
Qy	3952	TTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAGGACTTGGCGTCCAAGCTCTCAGTGA	4011
Db	3736	TTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAGGACTTGGCGTCCAAGCTCTCAGTGA	3795
Qy	4012	TCCCTCAAGATCCAGTGTCTCTCAGGAACCATCAGATTCAACTTAGATCCTTTGACC	4071
Db	3796	TCCCTCAAGATCCAGTGTCTCTCAGGAACCATCAGATTCAACTTAGATCCTTTGACC	3855
Qy	4072	GTCACACTGACCAGCAGATCTGGGATGCCCTTGGAGAGGACATTCCTTGACC AAGGCCATCT	4131

Db	3856	GTCACTGACCAGCAGATCTGGGATGCCCTGGAGAGGACATTCCTGACCAAGGCCATCT	3915
QY	4132	CAAAAGTCCCCAAAAGCTGCATACAGATGTGGTGGAAAAACGGTGGAAACTTCTCTGTGG	4191
Db	3916	CAAAAGTCCCCAAAAGCTGCATACAGATGTGGTGGAAAAACGGTGGAAACTTCTCTGTGG	3975
QY	4192	GGGAGAGGCAGCTGCTCTGCATTGCCAGGGCTGTGCTTCGCAACTCCAAAGATCATCCTTA	4251
Db	3976	GGGAGAGGCAGCTGCTCTGCATTGCCAGGGCTGTGCTTCGCAACTCCAAAGATCATCCTTA	4035
QY	4252	TCGATGAAGCCACAGCCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATCC	4311
Db	4036	TCGATGAAGCCACAGCCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATCC	4095
QY	4312	GTGAAGCCTTCAGGGCTGCACCGTGTCTGCTATGGCAATGGGAAGGTGGTAGAATTTGATCGCGCGGAGG	4371
Db	4096	GTGAAGCCTTCAGGGCTGCACCGTGTCTGCTATGGCAATGGGAAGGTGGTAGAATTTGATCGCGCGGAGG	4155
QY	4372	ACTGTGACCACATCCTCGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGCGCGGAGG	4431
Db	4156	ACTGTGACCACATCCTCGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGCGCGGAGG	4215
QY	4432	TACTGCGGAAGAACCTGGGTCAATGTTCCAGCCCTCATGGCCACACGCCACTTCTTCAC	4491
Db	4216	TACTGCGGAAGAACCTGGGTCAATGTTCCAGCCCTCATGGCCACACGCCACTTCTTCAC	4275
QY	4492	TGAGATAAGGAGATGTGGAGACTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTCCACACA	4551
Db	4276	TGAGATAAGGAGATGTGGAGACTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTCCACACA	4335
QY	4552	GGTGCAGCTTCGAGGCCCCACAGTCTGCGACCTTCTTGTTGGAGATCAGAACTTCTCCTG	4611
Db	4336	GGTGCAGCTTCGAGGCCCCACAGTCTGCGACCTTCTTGTTGGAGATCAGAACTTCTCCTG	4395
QY	4612	GAAGCAGGGTAAATGTAGGGGGGTGGGATTGCTGGATGGAAACCCTGGAATAGGCTA	4671
Db	4396	GAAGCAGGGTAAATGTAGGGGGGTGGGATTGCTGGATGGAAACCCTGGAATAGGCTA	4455
QY	4672	CTTGATGGCTCTCAAGACCTTAGAACCCAGAACCATCTAAGACATGGGATTCAGTGATC	4731
Db	4456	CTTGATGGCTCTCAAGACCTTAGAACCCAGAACCATCTAAGACATGGGATTCAGTGATC	4515
QY	4732	ATGTGGTTCCTTTAACTTACATGCTGAATAATTTTATAATAAGTAAAAGCTTATAG	4791
Db	4516	ATGTGGTTCCTTTAACTTACATGCTGAATAATTTTATAATAAGTAAAAGCTTATAG	4575
QY	4792	TTTTCTGATCTGTGTAGAAAGTGTGCAATGCTGTACTGACTTTGFAAAATATAAAACT	4851
Db	4576	TTTTCTGATCTGTGTAGAAAGTGTGCAATGCTGTACTGACTTTGFAAAATATAAAACT	4635
QY	4852	AAG 4854	
Db	4636	AAG 4638	
RESULT 5			
AF411579			
LOCUS			
DEFINITION Homo sapiens ATP-binding cassette protein C11 isoform A (ABCC11)			
AF411579			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			

PUBMED	11688999																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
--------	----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 358 TACCTCACCGTGTATGGCTCACCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGAT 417

Qy 690 GAGAACACCATCCCTCCACTGTCACTGTCACTCCATGATGCCTCAGACAAAAATGTCCAAAGGCTT 749

Db 418 GAGAACACCATCCCTCCACTGTCACTGTCACTCCATGATGCCTCAGACAAAAATGTCCAAAGGCTT 477

Qy 750 CACCGCCTTTGGGAAGAAGTCTCAAGGCGAGGGATTGAAAAAGCTTCAGTGCCTCTG 809

Db 478 CACCGCCTTTGGGAAGAAGTCTCAAGGCGAGGGATTGAAAAAGCTTCAGTGCCTCTG 537

Qy 810 GTGATGCTGAGGTTCCAGAGAACAAAGGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTC 869

Db 538 GTGATGCTGAGGTTCCAGAGAACAAAGGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTC 597

Qy 870 TGCATTGCCAGTGTACTCGGGCCAATATTGATTTATACAAAGATCCTGGAATATTCAGAA 929

Db 598 TGCATTGCCAGTGTACTCGGGCCAATATTGATTTATACAAAGATCCTGGAATATTCAGAA 657

Qy 930 GAGCAGTTGGGGAATGTTGTCCATGGAGTGGGACTCTGCTTGTGCCCTTTTCTCTCCGAA 989

Db 658 GAGCAGTTGGGGAATGCTGTCCATGGAGTGGGACTCTGCTTGTGCCCTTTTCTCTCCGAA 717

Qy 990 TGTGTGAAGTCTCTGAGTTTCTCCTCCAGTTGGATCATCAACCAACGCACAGCCCATCAGG 1049

Db 718 TGCCTGAAGTCTCTGAGTTTCTCCTCCAGTTGGATCATCAACCAACGCACAGCCCATCAGG 777

Qy 1050 TTCCGAGCAGCTGTTTCCCTCCTTTGCCCTTTGAGAAAGCTCATCCAATTTAAGTCTGTAATA 1109

Db 778 TTCCGAGCAGCTGTTTCCCTCCTTTGCCCTTTGAGAAAGCTCATCCAATTTAAGTCTGTAATA 837

Qy 1110 CACATCACCTCAGGAGAGGCCATCAGCTTCTTCCACCGGTGATGTAAACTACCTGTTTGAA 1169

Db 838 CACATCACCTCAGGAGAGGCCATCAGCTTCTTCCACCGGTGATGTAAACTACCTGTTTGAA 897

Qy 1170 GGGGTGTGCTATGGACCCCTAGTACTGATCACCTGGCGATCGCTGGTCACTGCAAGCATT 1229

Db 898 GGGGTGTGCTATGGACCCCTAGTACTGATCACCTGGCGATCGCTGGTCACTGCAAGCATT 957

Qy 1230 TCTTCCTACTTTCATTTATGGATACACTGCATTTATGGCATCTTATGCTATCTCCTGGTT 1289

Db 958 TCTTCCTACTTTCATTTATGGATACACTGCATTTATGGCATCTTATGCTATCTCCTGGTT 1017

Qy 1290 TTCCGACTGGCGGTATTATGACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAG 1349

Db 1018 TTCCGACTGGCGGTATTATGACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAG 1077

Qy 1350 GTCAGCGACGAGCGCATCCGCTGACCAGTGAAGTTCTCACTTGCATTAAGCTGATTAA 1409

Db 1078 GTCAGCGACGAGCGCATCCGCTGACCAGTGAAGTTCTCACTTGCATTAAGCTGATTAA 1137

Qy 1410 ATGTACACATGGGAGAAACCATTTGCAAAAAATCATTGAAGACCTAAGAAGGAAGAAAGG 1469

Db 1138 ATGTACACATGGGAGAAACCATTTGCAGAAATCATTGAAGACCTAAGAAGGAAGAAAGG 1197

Qy 1470 AAACTATTGGAGAAAGTGGCGCTTGTCCAGAGCCTGACAAGTATAACCTTGTTCATCATC 1529

Db 1198 AAACTATTGGAGAAAGTGGCGCTTGTCCAGAGCCTGACAAGTATAACCTTGTTCATCATC 1257

Qy 1530 CCCACAGTGGCCACAGCGTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACA 1589

Db 1258 CCCACAGTGGCCACAGCGTCTGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACA 1317

Qy 1590 GCGTCAATGGCCTTCAGCATGCTGGCCTCCTTGAATCTCCTTCGGCTGTCACTTCTTT 1649

Db 1318 GCGTCAATGGCCTTCAGCATGCTGGCCTCCTTGAATCTCCTTCGGCTGTCACTTCTTT 1377

Qy 1650 GTGCCATTGCAAGTCAAAAGTCTCACGAAATCCAAAGTCTGCAGTGTAGGTTCAAGAAG 1709

Db 1378 GTGCCATTGCAAGTCAAAAGTCTCACGAAATCCAAAGTCTGCAGTGTAGGTTCAAGAAG 1437

Qy 1710 TTTTCCCTCCAGGAGAGCCCTGTTTCTATGTCCAGACATTTACAAGACCCCAAGCT 1769

Db 1438 TTTTCTCCAGGAGAGCCCTGTTTCTATGTCCAGACATTACAAAGACCCCAAGAGCT 1497

Qy 1770 CTGGTCTTTGAGGAGGCCACCTTGTTCATGGCAACAGACCTGTCCCAGGATCGTCAATGGG 1829

Db 1498 CTGGTCTTTGAGGAGGCCACCTTGTTCATGGCAACAGACCTGTCCCAGGATCGTCAATGGG 1557

Qy 1830 GCACTGGAGCTGGAGAGGAACGGGCATGCTTCTGAGGGGATGACCAGGCCTAGAGATGCC 1889

Db 1558 GCACTGGAGCTGGAGAGGAACGGGCATGCTTCTGAGGGGATGACCAGGCCTAGAGATGCC 1617

Qy 1890 CTCGGGCCAGAGGAAGAAGGAACAGCCTGGGCCCCAGAGTTGCACAAGATCAACCTGGTG 1949

Db 1618 CTCGGGCCAGAGGAAGAAGGAACAGCCTGGGCCCCAGAGTTGCACAAGATCAACCTGGTG 1677

Qy 1950 GTGTCCAAGGGGATGATGTTAGGGGTCTGCGGCAACACGGGGAGTGGTAAGAGCAGCCTG 2009

Db 1678 GTGTCCAAGGGGATGATGTTAGGGGTCTGCGGCAACACGGGGAGTGGTAAGAGCAGCCTG 1737

Qy 2010 TTGTCAGCCATCCTGGAGGAGATGCACCTTGCTCGAGGGTCTGGTGGGGTGCAGGAAAGC 2069

Db 1738 TTGTCAGCCATCCTGGAGGAGATGCACCTTGCTCGAGGGTCTGGTGGGGTGCAGGAAAGC 1797

Qy 2070 CTGGCCTATGTCCCCAGCAGGCCTGGATCGTCAGGGGAACATCAGGGAGAACATCCTTC 2129

Db 1798 CTGGCCTATGTCCCCAGCAGGCCTGGATCGTCAGCGGAACATCAGGGAGAACATCCTTC 1857

Qy 2130 ATGGGAGGCGCATATGACAAGGCCCGCATACCTCCAGGTGCTCCACTGCTGCTCCCTGAAAT 2189

Db 1858 ATGGGAGGCGCATATGACAAGGCCCGCATACCTCCAGGTGCTCCACTGCTGCTCCCTGAAAT 1917

Qy 2190 CGGACCTGGAACCTTCTGCCCTTTGGAGACATGACAGAGATTGGAGAGCGGGGCTCAAC 2249

Db 1918 CGGACCTGGAACCTTCTGCCCTTTGGAGACATGACAGAGATTGGAGAGCGGGGCTCAAC 1977

Qy 2250 CTCTCTGGGGGCAGAAACAGAGGATCAGCCTGGCCCGCGCGTCTATTCCGACCGTCAAG 2309

Db 1978 CTCTCTGGGGGCAGAAACAGAGGATCAGCCTGGCCCGCGCGTCTATTCCGACCGTCAAG 2037

Qy 2310 ATCTACCTGCTGGACGACCCCTGTCTGTGTGGAGCCCACTGGGGAAGCAGATTTT 2369

Db 2038 ATCTACCTGCTGGACGACCCCTGTCTGTGTGGAGCCCACTGGGGAAGCAGATTTT 2097

Qy 2370 GAGGAGTGCAATTAAGAAACACTCAGGGGGAAGACGCTCGTCTGGTACCCACAGCTG 2429

Db 2098 GAGGAGTGCAATTAAGAAACACTCAGGGGGAAGACGCTCGTCTGGTACCCACAGCTG 2157

Qy 2430 CAGTACTTAGAATTTTGTGGCCAGATCAJTTTGTGAAAAATGGAAAAATCTGTGAAAAAT 2489

Db 2158 CAGTACTTAGAATTTTGTGGCCAGATCATTTTGTGAAAAATGGAAAAATCTGTGAAAAAT 2217

Qy 2490 GGAACTCACAGTGAGTTAATGCAGAAAAAGGGGAATATATGCCCCAACCTTATCCAGAAGATG 2549

Db 2218 GGAACTCACAGTGAGTTAATGCAGAAAAAGGGGAATATATGCCCCAACCTTATCCAGAAGATG 2277

Qy 2550 CACAAGGAAGCCACTTCGGACATGTTGCAGGACACAGCAAAAGATAGCAGAGAAAGCCAAAG 2609

Db 2278 CACAAGGAAGCCACTTCGGACATGCTGCAGGACACAGCAAAAGATAGCAGAGAAAGCCAAAG 2337

Qy 2610 GTAGAAAGTCAGGCTCTGGCCACCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCCG 2669

Db 2338 GTAGAAAGTCAGGCTCTGGCCACCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCCG 2397

Qy 2670 GAGCATCAGCTCACACAGGAGGAGAGATGGAAGAAGGCTCCTTGAGTTGGAGGTTCTAC 2729

Db 2398 GAGCATCAGCTCACACAGGAGGAGAGATGGAAGAAGGCTCCTTGAGTTGGAGGTTCTAC 2457

Qy 2730 CACCACTACATCCAGCAGCTGGAGGTTACATGGTCTCTTGATATAATTTTCTTCTCGTG 2789

Db 2458 CACCACTACATCCAGCAGCTGGAGGTTACATGGTCTCTTGATATAATTTTCTTCTCGTG 2517

Qy 2790 GTGCTGATCGTCTTAAACGATCTTCAGCTTCTGGTGGCTGAGCTACTGTTGGAGCAG 2849

Db 2518 GTGCTGATCGTCTTAAACGATCTTCAGCTTCTGGTGGCTGAGCTACTGTTGGAGCAG 2577

QY	2850	GGCTCGGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATT	2909
Db	2578	GGCTCGGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATGGCAGACCTGGGCAACATT	2637
QY	2910	GCAGACAATCCTCAACTGTCCTTCTACCAAGCTGGTGACGGGCTCAACGCCCTGCTCCTC	2969
Db	2638	GCAGACAATCCTCAACTGTCCTTCTACCAAGCTGGTGACGGGCTCAACGCCCTGCTCCTC	2697
QY	2970	ATCTGTGTGGGGGTCTGCTCCTCAGGGATTTCACCAAAAGTCAACGAGGAAGGCATCCACG	3029
Db	2698	ATCTGTGTGGGGGTCTGCTCCTCAGGGATTTCACCAAAAGTCAACGAGGAAGGCATCCACG	2757
QY	3030	GCCCTGCACAACAAGCTCTTCAACAAGGTTTCCCGCTGCCCATGAGTTTCTTTGACACC	3089
Db	2758	GCCCTGCACAACAAGCTCTTCAACAAGGTTTCCCGCTGCCCATGAGTTTCTTTGACACC	2817
QY	3090	ATCCCAATAGCGGCTTTTGAACGTCTTCGAGGGGACTTGGAAACAGCTGGACCAGCTC	3149
Db	2818	ATCCCAATAGCGGCTTTTGAACGTCTTCGAGGGGACTTGGAAACAGCTGGACCAGCTC	2877
QY	3150	TTGCCCATCTTTTCAGAGCAGTTCCCTGCTCCTTAATGGTGATCCCGTCTCTGTTG	3209
Db	2878	TTGCCCATCTTTTCAGAGCAGTTCCCTGCTCCTTAATGGTGATCCCGTCTCTGTTG	2937
QY	3210	ATTGTCAGTGTGCTCTCCATATATCCTGTAAATGGGAGCCATAATCATGGTTATTGGC	3269
Db	2938	ATTGTCAGTGTGCTCTCCATATATCCTGTAAATGGGAGCCATAATCATGGTTATTGGC	2997
QY	3270	TTCATTTATTATATGATTCAAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTAT	3329
Db	2998	TTCATTTATTATATGATTCAAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTAT	3057
QY	3330	AGCCGGTCTCCTTTATTCTCCACATCCTCAATCTCTGCAAGGCCTGAGTCCATCCAT	3389
Db	3058	AGCCGGTCTCCTTTANTCTCCACATCCTCAATCTCTGCAAGGCCTGAGTCCATCCAT	3117
QY	3390	GTCTATGGAAAACCTGAAGACTTCATCAGCCAGTTTAAAGAGGCTGACTGCGCAGAAT	3449
Db	3118	GTCTATGGAAAACCTGAAGACTTCATCAGCCAGTTTAAAGAGGCTGACTGCGCAGAAT	3177
QY	3450	AACTACCTGCTGTGTTCTATCTTCCACACGATGGATGGCATTGAGGCTGGAGATCATG	3509
Db	3178	AACTAGCTGCTGTGTTCTATCTTCCACACGATGGATGGCATTGAGGCTGGAGATCATG	3237
QY	3510	ACCAACCTTGTGACCTTGGCTGTTGCCCTGTTCGTGGCTTTTGGCATTTCCCTCCACCCCC	3569
Db	3238	ACCAACCTTGTGACCTTGGCTGTTGCCCTGTTCGTGGCTTTTGGCATTTCCCTCCACCCCC	3297
QY	3570	TACTCCTTTAAAGTCATGGCTGTCAACATCGTGCAGCTGGCGTCCAGCTTCCAGGCC	3629
Db	3298	TACTCCTTTAAAGTCATGGCTGTCAACATCGTGCAGCTGGCGTCCAGCTTCCAGGCC	3357
QY	3630	ACTGCCCGGATTGGCTTGGAGACAGAGGCACAGTTACAGGCTGTAGAGAGGATACTGCAG	3689
Db	3358	ACTGCCCGGATTGGCTTGGAGACAGAGGCACAGTTACAGGCTGTAGAGAGGATACTGCAG	3417
QY	3690	TACATGAAGATGTGTGTCGGAAGCTCCTTACACATGGAAGGCACAAGTTGTCCCCAG	3749
Db	3418	TACATGAAGATGTGTGTCGGAAGCTCCTTACACATGGAAGGCACAAGTTGTCCCCAG	3477
QY	3750	GGGTGGCCACAGCATGGGGAATCATATTTCAGGATTATCACATGAATAACAGAGACAAC	3809
Db	3478	GGGTGGCCACAGCATGGGGAATCATATTTCAGGATTATCACATGAATAACAGAGACAAC	3537
QY	3810	ACACCCACCGTGTCTCACGGCATCAACCTGACCATCCGCGGCCACGAAGTGGTGGGCATC	3869
Db	3538	ACACCCACCGTGTCTCACGGCATCAACCTGACCATCCGCGGCCACGAAGTGGTGGGCATC	3597
QY	3870	GTGGGAAGGACGGGCTCTGGGAAGTCTCCTTGGGCATGGCTCTCTCCGCCCTGGTGGAG	3929
Db	3598	GTGGGAAGGACGGGCTCTGGGAAGTCTCCTTGGGCATGGCTCTCTCCGCCCTGGTGGAG	3657

QY	3930	CCCATGGCAGCCGGGATTCTCATTTGACGGCGTGGACATTTGCAGCATCGGCCTGGAGGAC	3989
Db	3658	CCCATGGCAGCCGGGATTCTCATTTGACGGCGTGGACATTTGCAGCATCGGCCTGGAGGAC	3717
QY	3990	TTGCGGTCCAAGCTCTCAGTGATCCCTCAAGATCCAGTGCTGCTCTCAGGAACCATCAGA	4049
Db	3718	TTGCGGTCCAAGCTCTCAGTGATCCCTCAAGATCCAGTGCTGCTCTCAGGAACCATCAGA	3777
QY	4050	TTCAACCTAGATCCCTTTGACCCGTCACACTGACCAGCAGATCTGGGATGCCCTTGGAGAGG	4109
Db	3778	TTCAACCTAGATCCCTTTGACCCGTCACACTGACCAGCAGATCTGGGATGCCCTTGGAGAGG	3837
QY	4110	ACATTCTCTGACCAAGGCCATCTCAAAGTTTCCCCAAAAGCTGCATACAGATGTGGTGGAA	4169
Db	3838	ACATTCTCTGACCAAGGCC-----ATCATCCTTATCGATGAAGCCACAGCCTCCATTGACATGGAGACAGAC	3855
QY	4170	AACGGTGGAAACTTCTCTGTGGGGGAGAGCAGCTGCTCTGCATTGCCAGGGCTGTGCTT	4229
Db	3856	-----ATCATCCTTATCGATGAAGCCACAGCCTCCATTGACATGGAGACAGAC	3855
QY	4230	CGAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCTCCATTGACATGGAGACAGAC	4289
Db	3856	-----ATCATCCTTATCGATGAAGCCACAGCCTCCATTGACATGGAGACAGAC	3903
QY	4290	ACCCTGATCCAGCGCACAAATCCGTGAAGCCTTCCAGGGCTGCACCCTGCTCGTCATTGCC	4349
Db	3904	ACCCTGATCCAGCGCACAAATCCGTGAAGCCTTCCAGGGCTGCACCCTGCTCGTCATTGCC	3963
QY	4350	CACCGTGTCAACCACTGTGCTGAACCTGTGACCACACATCCTGGTTATGGGCAATGGGAAGGTG	4409
Db	3964	CACCGTGTCAACCACTGTGCTGAACCTGTGACCACACATCCTGGTTATGGGCAATGGGAAGGTG	4023
QY	4410	GTAGAAATTGATCGGCCGGAGGTACTCGCGGAAGAAGCCCTGGGTCATTGTTCCGACGCCCTC	4469
Db	4024	GTAGAAATTGATCGGCCGGAGGTACTCGCGGAAGAAGCCCTGGGTCATTGTTCCGACGCCCTC	4083
QY	4470	ATGGCCACAGCCACTTCTTCACTGAGATAAGGAGATGTGGAGACTTCATGGAGGCTGGCA	4529
Db	4084	ATGGCCACAGCCACTTCTTCACTGAGATAAGGAGATGTGGAGACTTCATGGAGGCTGGCA	4143
QY	4530	GCTGAGCTCAGAGGTTTACACACAGGTGCAGCTTCGAGGCCACAGTCTCGACCTTCTTGT	4589
Db	4144	GCTGAGCTCAGAGGTTTACACACAGGTGCAGCTTCGAGGCCACAGTCTCGACCTTCTTGT	4203
QY	4590	TTGGAGATCAGAACTTCTCCTGGAAGCAGGGGTAAATGTAGGGGGTGGGATTTGCTGG	4649
Db	4204	TTGGAGATCAGAACTTCTCCTGGAAGCAGGGGTAAATGTAGGGGGTGGGATTTGCTGG	4263
QY	4650	ATGGAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCAAGAACCATC	4709
Db	4264	ATGGAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCAAGAACCATC	4323
QY	4710	TAAGACATGGGATTTCAGTGATCATGTGTTCTCCTTTTAACTTACATGCTGAATAATTTT	4769
Db	4324	TAAGACATGGGATTTCAGTGATCATGTGTTCTCCTTTTAACTTACATGCTGAATAATTTT	4383
QY	4770	ATAATAAGGTAAAAGCTTATAGTTTCTGATCTGTGTTAGAAAGTGTGCAAAATGCTGTAC	4829
Db	4384	ATAATAAGGTAAAAGCTTATAGTTTCTGATCTGTGTTAGAAAGTGTGCAAAATGCTGTAC	4443
QY	4830	TGACTTTGTAAAAATATAAAACTAAGGAAAACTC 4862	
Db	4444	TGACTTTGTAAAAATATAAAACTAAGGAAAACTC 4476	

RESULT 6					
AX135184					
LOCUS	AX135184	4149 bp	DNA	linear	PAT 29-MAY-2001
DEFINITION	Sequence 17 from Patent WO0132706.				
ACCESSION	AX135184				
VERSION	AX135184.1	GI:14271537			
KEYWORDS					
SOURCE	human.				

ORGANISM	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
REFERENCE	1 (bases 1 to 4149)									
AUTHORS	Harras,M., Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.									
TITLE	Novel human transporter proteins and polynucleotides encoding the same									
JOURNAL	Patent: WO 0132706-A 17 10-MAY-2001;									
FEATURES	Lexicon Genetics Incorporated (US)									
source	Location/Qualifiers									
	1..4149									
	/organism="Homo sapiens"									
	/db_xref="taxon:9606"									
BASE COUNT	982 a	1080 c	1091 g	996 t						
ORIGIN										
	Query Match 85.2%; Score 4142.6; DB 6; Length 4149;									
	Best Local Similarity 99.9%; Pred. No. 0;									
	Matches 4145; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
Qy	351	ATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	410							
Db	1	ATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	60							
Qy	411	ATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT	470							
Db	61	ATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT	120							
Qy	471	GGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGAGGCGAGCTGTCCCACCG	530							
Db	121	GGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGAGGCGAGCTGTCCCACCG	180							
Qy	531	TGGGGAGATATGATGCTGCCTTGAGAACCATGATTCCTTCCGTCCCAAGCGAGGTTT	590							
Db	181	TGGGGAGATATGATGCTGCCTTGAGAACCATGATTCCTTCCGTCCCAAGCGAGGTTT	240							
Qy	591	CCTGCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTCTACCTCACCGTGTCATGGCTC	650							
Db	241	CCTGCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTCTACCTCACCGTGTCATGGCTC	300							
Qy	651	ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTG	710							
Db	301	ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTG	360							
Qy	711	TCAGTCCATGATGCCTCAGACAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA	770							
Db	361	TCAGTCCATGATGCCTCAGACAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA	420							
Qy	771	GTCTCAAGCGGAGGGATTGAAAAAGCTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGA	830							
Db	421	GTCTCAAGCGGAGGGATTGAAAAAGCTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGA	480							
Qy	831	ACAAGGTTGATTTTCGATGCACCTCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGGG	890							
Db	481	ACAAGGTTGATTTTCGATGCACCTCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGGG	540							
Qy	891	CCAATATTGATTATACCAAAGATCCTGGAATATTCAGAAAGACAGTTGGGAATGTTGTC	950							
Db	541	CCAATATTGATTATACCAAAGATCCTGGAATATTCAGAAAGACAGTTGGGAATGTTGTC	600							
Qy	951	CATGAGTGGGACTCTGCTTTGGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC	1010							
Db	601	CATGAGTGGGACTCTGCTTTGGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC	660							
Qy	1011	TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCC	1070							
Db	661	TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCC	720							
Qy	1071	TTTGCCCTTTGAGAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC	1130							
Db	721	TTTGCCCTTTGAGAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC	780							

Db 1861 TTTGGAGACATGACAGAGATTGGAGCGGGGCTCAACCTCTCTGGGGGCAGAAACAG 1920
QY 2271 AGGATCAGCCTGGCCCGCCGCTCTATTCCGACCGTCAAGATCTACCTGCTGGACGACCCC 2330
Db 1921 AGGATCAGCCTGGCCCGCCGCTCTATTCCGACCGTCAAGATCTACCTGCTGGACGACCCC 1980
QY 2331 CTGTCTGCTGTGGACGCCACAGTGGGGAAGACATTTTGGAGAGTGCATTAAGAAGACA 2390
Db 1981 CTGTCTGCTGTGGACGCCACAGTGGGGAAGACATTTTGGAGAGTGCATTAAGAAGACA 2040
QY 2391 CTCAGGGGAAGACGGTCTCTCTGGTGACCCACAGCTGCAGTACTTGAATTTTGTGGC 2450
Db 2041 CTCAGGGGAAGACGGTCTCTCTGGTGACCCACAGCTGCAGTACTTGAATTTTGTGGC 2100
QY 2451 CAGATCATTTTGTGGAATAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAAGTAAATG 2510
Db 2101 CAGATCATTTTGTGGAATAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAAGTAAATG 2160
QY 2511 CAGAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAGCCACTTCGGAC 2570
Db 2161 CAGAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAGCCACTTCGGAC 2220
QY 2571 ATGTTGCAGGACACAGCAAGATAGCAGAGAGCCAAAGGTAGAAAGTCAGGCTCTGGCC 2630
Db 2221 ATGTTGCAGGACACAGCAAGATAGCAGAGAGCCAAAGGTAGAAAGTCAGGCTCTGGCC 2280
QY 2631 ACCTCCCTGGAAGAGTCTCTCAACGGGAAATGCTGTGCCGGAGCATCAGTCAACACAGGAG 2690
Db 2281 ACCTCCCTGGAAGAGTCTCTCAACGGGAAATGCTGTGCCGGAGCATCAGTCAACACAGGAG 2340
QY 2691 GAGGAGATGGAAGAGGCTCTCTTGAGTTGGAGGTCTACCACCACCTACATCCAGGCAGCT 2750
Db 2341 GAGGAGATGGAAGAGGCTCTCTTGAGTTGGAGGTCTACCACCACCTACATCCAGGCAGCT 2400
QY 2751 GGAGGTTACATGFTCTCTTGCAATAATTTCTCTCGTGGTCTGATCGTCTCTTAACG 2810
Db 2401 GGAGGTTACATGFTCTCTTGCAATAATTTCTCTCGTGGTCTGATCGTCTCTTAACG 2460
QY 2811 ATCTTCAGCTTCGTGGCTGAGCTACTGFTGGAGCAGGGCTCGGGGACCAATAGCAGC 2870
Db 2461 ATCTTCAGCTTCGTGGCTGAGCTACTGFTGGAGCAGGGCTCGGGGACCAATAGCAGC 2520
QY 2871 CGAGAGAGCAATGGAACCATGGCAGACCTGSGCAACATTGCAGACAATCCTCAACTGTCC 2930
Db 2521 CGAGAGAGCAATGGAACCATGGCAGACCTGSGCAACATTGCAGACAATCCTCAACTGTCC 2580
QY 2931 TTCTACCAGCTGGTGACGGGCTCAACGGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCC 2990
Db 2581 TTCTACCAGCTGGTGACGGGCTCAACGGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCC 2640
QY 2991 TCAGGGATTTTCACCAAAAGTCACGAGGAAGGCATCCACGGCCCTGCACAACAAGCTCTTC 3050
Db 2641 TCAGGGATTTTCACCAAAAGTCACGAGGAAGGCATCCACGGCCCTGCACAACAAGCTCTTC 2700
QY 3051 AACAAAGTTTTCGGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGGCCGGCTTTTG 3110
Db 2701 AACAAAGTTTTCGGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGGCCGGCTTTTG 2760
QY 3111 AACTGCTTCGACGGGACTTGGAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAGCAG 3170
Db 2761 AACTGCTTCGACGGGACTTGGAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAGCAG 2820
QY 3171 TTCCTGGTCCTTAATGGTGATCGCCGCTCTGTTGATTGTCAAGTGTGCTCTCCA 3230
Db 2821 TTCCTGGTCCTTAATGGTGATCGCCGCTCCTGTTGATTGTCAAGTGTGCTCTCCA 2880
QY 3231 TATATCCTGTTAATGGGAGCCATAATCATGGTTATTGGCTTCATTTATATATGATGTTTC 3290
Db 2881 TATATCCTGTTAATGGGAGCCATAATCATGGTTATTGGCTTCATTTATATATGATGTTTC 2940
QY 3291 AAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATCTCC 3350
Db 4021 AACTGTGACCACATCCTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCGGAG 4080

Db 2941 AAGAAAGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTTCTCC 3000
QY 3351 CACATCCTCAATTTCTCTGAAGGCCCTGAGTCCATCCATGTCTATGGAANAACCTGAAGAC 3410
Db 3001 CACATCCTCAATTTCTCTGAAGGCCCTGAGTCCATCCATGTCTATGGAANAACCTGAAGAC 3060
QY 3411 TTCAATCAGCCAGTTTAAGAGGCTGACTGATCGCGCAGAATAACTACCTGCTTTGTTCTA 3470
Db 3061 TTCAATCAGCCAGTTTAAGAGGCTGACTGATCGCGCAGAATAACTACCTGCTTTGTTCTA 3120
QY 3471 TCTTCCACACGATGGATGGCATTTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT 3530
Db 3121 TCTTCCACACGATGGATGGCATTTGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT 3180
QY 3531 GTTGGCCCTGTCTGGGCTTTTGGCATTTTCCACCCCTACTCCTTTAAAGTCATGGCT 3590
Db 3181 GTTGGCCCTGTCTGGGCTTTTGGCATTTTCCACCCCTACTCCTTTAAAGTCATGGCT 3240
QY 3591 GTCAACATCGTCTGCAGCTGGCTCCAGTTCCAGGCCACTGCCGGATTGGCTTGGAG 3650
Db 3241 GTCAACATCGTCTGCAGCTGGCTCCAGTTCCAGGCCACTGCCGGATTGGCTTGGAG 3300
QY 3651 ACAGAGGCACAGTTTACGGCTGTAGAGAGATACTGCAGTACATGAAGATGTGTCTCG 3710
Db 3301 ACAGAGGCACAGTTTACGGCTGTAGAGAGATACTGCAGTACATGAAGATGTGTCTCG 3360
QY 3711 GAAGCTCCTTTACACATGGAAGGCACAAAGTTGTCCCCAGGGTGGCCACAGCATGGGAA 3770
Db 3361 GAAGCTCCTTTACACATGGAAGGCACAAAGTTGTCCCCAGGGTGGCCACAGCATGGGAA 3420
QY 3771 ATCATATTTTCAAGATTATCACATGAATACAGAGACAACACACCCACCGTCTTCACGGC 3830
Db 3421 ATCATATTTTCAAGATTATCACATGAATACAGAGACAACACACCCACCGTCTTCACGGC 3480
QY 3831 ATCAACCTGACCATCCGGGGCCACGAAGTGGTGGGCATCGTGGGAAGGACGGGCTCTGGG 3890
Db 3481 ATCAACCTGACCATCCGGGGCCACGAAGTGGTGGGCATCGTGGGAAGGACGGGCTCTGGG 3540
QY 3891 AAGTCTCCTTGGGCATGGCTCTCTTCCGCTGGTGGAGCCCATGGCAGGCCGGATTCTC 3950
Db 3541 AAGTCTCCTTGGGCATGGCTCTCTTCCGCTGGTGGAGCCCATGGCAGGCCGGATTCTC 3600
QY 3951 ATTGACGGCTGGACATTTGCAGCATCGGCTGGAGGACTTGCGGTCCAAGCTCTCAGTG 4010
Db 3601 ATTGACGGCTGGACATTTGCAGCATCGGCTGGAGGACTTGCGGTCCAAGCTCTCAGTG 3660
QY 4011 ATCCCTCAAGATCCAGTGTCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTTGAC 4070
Db 3661 ATCCCTCAAGATCCAGTGTCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTTGAC 3720
QY 4071 CGTCACACTGACCAGCAGATCTGGGATGCTTGGAGAGGACATTCCTGACCAAGGCCATC 4130
Db 3721 CGTCACACTGACCAGCAGATCTGGGATGCTTGGAGAGGACATTCCTGACCAAGGCCATC 3780
QY 4131 TCAAAGTTCCCCAAAAGCTGCATACAGATGTGGTGGAAAACGGTGGAAAACCTTCTCTGTG 4190
Db 3781 TCAAAGTTCCCCAAAAGCTGCATACAGATGTGGTGGAAAACGGTGGAAAACCTTCTCTGTG 3840
QY 4191 GGGAGAGGAGCTGCTCTGCATTTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCCTT 4250
Db 3841 GGGAGAGGAGCTGCTCTGCATTTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCCTT 3900
QY 4251 ATCGATGAAGCCACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATC 4310
Db 3901 ATCGATGAAGCCACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATC 3960
QY 4311 CGTGAAGCCTTCCAGGGCTGCACCGTGTCTCATTTGCCACCCGTTGCACCACTGTGCTG 4370
Db 3961 CGTGAAGCCTTCCAGGGCTGCACCGTGTCTCATTTGCCACCCGTTGCACCACTGTGCTG 4020
QY 4371 AACTGTGACCACATCCTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCGGAG 4430
Db 4021 AACTGTGACCACATCCTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCGGAG 4080

Qy 4431 GTACTGCGGAAGAAGCCTGGGTCAATTGTTGCGAGCCCTCATGGCCACAGCCACTTCTTCA 4490
|||||
Db 4081 GTACTGCGGAAGAAGCCTGGGTCAATTGTTGCGAGCCCTCATGGCCACAGCCACTTCTTCA 4140
Qy 4491 CTGAGATAA 4499
Db 4141 CTGAGATAA 4149

RESULT 7
AX135192
LOCUS AX135192 5020 bp DNA linear PAT 29-MAY-2001
DEFINITION Sequence 25 from Patent WO0132706.
ACCESSION AX135192
VERSION AX135192.1 GI:14271541
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5020)
AUTHORS Harras,M., Donoho,G., Turner,C.A., Nehls,M., Friedrich,G.,
Zambrowicz,B. and Sands,A.T.
TITLE Novel human transporter proteins and polynucleotides encoding the
same
JOURNAL Patent: WO 0132706-A 25 10-MAY-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source Location/Qualifiers
1..5020
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1240 a 1249 c 1329 g 1202 t
ORIGIN

Query Match 85.0%; Score 4133.6; DB 6; Length 5020;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 4529; Conservative 0; Mismatches 4; Indels 363; Gaps 3;

Qy 330 AGAATTGGCAGGAAC TGAAATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCT 389
|||||
Db 99 AGAATTGGCAGGAAC TGAATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCT 158

Qy 390 GGTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTATTAT 449
|||||
Db 159 GGTGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTATTAT 218

Qy 450 AAAACCTATACTCTCCAAAGATGGCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCA 509
|||||
Db 219 AAAACCTATACTCTCCAAAGATGGCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCA 278

Qy 510 GGGAGGGCAGCTGTCCACCGTGGGGGAAGTATGATGCTGCCTTGAGAACCATGATTCCC 569
|||||
Db 279 GGGAGGGCAGCTGTCCACCGTGGGGGAAGTATGATGCTGCCTTGAGAACCATGATTCCC 338

Qy 570 TTCCGTCCCCAAGCCGAGGTTTCCTGCCCCCCAGCCCTGGACAATGCTGGCCTGTTCTCC 629
|||||
Db 339 TTCCGTCCCCAAGCCGAGGTTTCCTGCCCCCCAGCCCTGGACAATGCTGGCCTGTTCTCC 398

Qy 630 TACCTCACCGTGTCA TGGCTCACCCCGCTCATGATCCAAAAGCTTACGGAGTCGCTTAGAT 689
|||||
Db 399 TACCTCACCGTGTCA TGGCTCACCCCGCTCATGATCCAAAAGCTTACGGAGTCGCTTAGAT 458

Qy 690 GAGAACACCATCCCTCCACTGTGAGTCCATGATGCCCTCAGACAAAAATGTCCAAAAGCCTT 749
|||||
Db 459 GAGAACACCATCCCTCCACTGTGAGTCCATGATGCCCTCAGACAAAAATGTCCAAAAGCCTT 518

Qy 750 CACCGCCTTTGGGAAGAAAGTCTCAAGGCGAGGGATTGAAAAAGCTTCAGTGCCTCTG 809
|||||
Db 519 CACCGCCTTTGGGAAGAAAGTCTCAAGGCGAGGGATTGAAAAAGCTTCAGTGCCTCTG 578

Qy 810 GTGATGCTGAGGTTCCAGAGAACAAAGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTC 869
|||||

Db 579 GTGATGCTGAGGTTCCAGAGAACAAAGGTTGATTTTCGATGCACCTTCTGGGCATCTTGCTTC 538
Qy 870 TGCATTGCCAGTGTACTCGGGCCA----- 893
|||||
Db 639 TGCATTGCCAGTGTACTCGGGCCAGACTGCACCAATGGGGAGGTGGGAGTCTTCCCAGA 598
Qy 894 -----ATATTGATTATACCAAAGAT 913
|||||
Db 699 AGAAAACTGAGGCAC TGGTACCAGCTACGGGGAGGTGAAGATATTGATTATACCAAAGAT 758

Qy 914 CCTGGAATATT CAGAAGAGCAGTTGGGGAATGTTGCCATGGAGTGGGACTCTGCTTGC 973
Db 759 CCTGGAATATT CAGAAGAGCAGTTGGGGAATGTTGCCATGGAGTGGGACTCTGCTTGC 818
Qy 974 CCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTCTCCTCCAGTTGGATCATCAACCA 1033
Db 819 CCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTCTCCTCCAGTTGGATCATCAACCA 878

Qy 1034 ACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCCTTTGCCCTTTGAGAAAGTCAATCCA 1093
|||||
Db 879 ACGCACAGCCATCAGGTTCCAAGCAGCTGTTTCCTCCTTTGCCCTTTGAGAAAGTCAATCCA 938

Qy 1094 ATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCCCATCAGCTTCTTACCGGTGATGT 1153
Db 939 ATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCCCATCAGCTTCTTACCGGTGATGT 998

Qy 1154 AAAC TACCTGTTTGAAGGGTGTGCTATGGACCCCTAGTACTGATCACCTGCGCATCGCT 1213
Db 999 AAAC TACCTGTTTGAAGGGTGTGCTATGGACCCCTAGTACTGATCACCTGCGCATCGCT 1058

Qy 1214 GGTCATCTGCAGCATTTCTTCCTACTTCA TATTGGATACACTGCATTTATTGCCATCTT 1273
Db 1059 GGTCATCTGCAGCATTTCTTCCTACTTCA TATTGGATACACTGCATTTATTGCCATCTT 1118

Qy 1274 ATGCTATCTCCTGGTTTTCCCAC TGGCGGTATTTCATGACAAAGAA TGGCTGTGAAGGCTCA 1333
Db 1119 ATGCTATCTCCTGGTTTTCCCAC TGGAGGTATTTCATGACAAAGAA TGGCTGTGAAGGCTCA 1178

Qy 1334 GCATCACACATCTGAGGTCAGGACCAGCGCATCCG TGTGACCAGTGAAGTTCTCACTTG 1393
Db 1179 GCATCACACATCTGAGGTCAGGACCAGCGCATCCG TGTGACCAGTGAAGTTCTCACTTG 1238

Qy 1394 CATTAAGCTGATTTAAATGTACACATGGGAGAA ACCATTTCGAAAAATCATTTGAAGACCT 1453
Db 1239 CATTAAGCTGATTTAAATGTACACATGGGAGAA ACCATTTCGAAAAATCATTTGAAGACCT 1298

Qy 1454 AAGAAGGAAGGAAAGAACTATTGGAGAA GTGCGGGCTTGTCCAGAGCCTGACAAGTAT 1513
Db 1299 AAGAAGGAAGGAAAGAACTATTGGAGAA GTGCGGGCTTGTCCAGAGCCTGACAAGTAT 1358

Qy 1514 AACCTTGTTCATCATCCCACAGTGGCCACAGCGGTCTGGGTTCTCATCCACACATCCTT 1573
Db 1359 AACCTTGTTCATCATCCCACAGTGGCCACAGCGGTCTGGGTTCTCATCCACACATCCTT 1418

Qy 1574 AAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATGCTGGCCTCCTTGAATCTCCTTCG 1633
Db 1419 AAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATGCTGGCCTCCTTGAATCTCCTTCG 1478

Qy 1634 GCTGTCAGTGTCTTTGTGCCTATTGCAGTCAAAGGTCTCACGAATTCCAAAGTCTGCAGT 1693
Db 1479 GCTGTCAGTGTCTTTGTGCCTATTGCAGTCAAAGGTCTCACGAATTCCAAAGTCTGCAGT 1538

Qy 1694 GATGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCTGTTTCTATGTCCAGACATTACA 1753
Db 1539 GATGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCTGTTTCTATGTCCAGACATTACA 1598

Qy 1754 AGACCC CAGCAAAAGCTCTGGTCTTTGAGGAGGCCACCTTGTTCATGGCAACAGACCTGTCC 1813
Db 1599 AGACCC CAGCAAAAGCTCTGGTCTTTGAGGAGGCCACCTTGTTCATGGCAACAGACCTGTCC 1658

Qy 1814 CGGGATCGTCAATGGGGCACTGGAGCTGGAGAGGAAACGGGCATGCTTCTGAGGGGATGAC 1873
Db 1659 CGGGATCGTCAATGGGGCACTGGAGCTGGAGAGGAAACGGGCATGCTTCTGAGGGGATGAC 1718

Db	 3879	CAGGGTGGCCACAGCATGGGAAATCATATTTCAGGATTATACATGAAATACAGAGAC	3938
Qy	 3807	AACACACCCACCGTGTCTTACGGCATCAACCTGACCATCCGCGGCCACGAAGTGGTGGC	3866
Db	 3939	AACACACCCACCGTGTCTTACGGCATCAACCTGACCATCCGCGGCCACGAAGTGGTGGC	3998
Qy	 3867	ATCGTGGGAAGACGGGCTCTGGAAGTCCCTCTTGGGCATGGCTCTCTTCCGCCTGGTG	3926
Db	 3999	ATCGTGGGAAGACGGGCTCTGGAAGTCCCTCTTGGGCATGGCTCTCTTCCGCCTGGTG	4058
Qy	 3927	GAGCCCATGGCAGGCCGGATTCTTCATTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAG	3986
Db	 4059	GAGCCCATGGCAGGCCGGATTCTTCATTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAG	4118
Qy	 3987	GACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAGATCCAGTCTGCTCTCAGGAACCATC	4046
Db	 4119	GACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAGATCCAGTCTGCTCTCAGGAACCATC	4178
Qy	 4047	AGATTCAACCTAGATCCCTTTGACCGTCACTGACCCAGCAGATCTGGGATGCCTTGGAG	4106
Db	 4179	AGATTCAACCTAGATCCCTTTGACCGTCACTGACCCAGCAGATCTGGGATGCCTTGGAG	4238
Qy	 4107	AGGACATTCCTGACCAAGGCCATCTCAAAAGTTCCCCAAAAAGCTGCATACAGATGTGGTG	4166
Db	 4239	AGGACATTCCTGACCAAGGCCATCTCAAAAGTTCCCCAAAAAGCTGCATACAGATGTGGTG	4298
Qy	 4167	GAARACGGTGGAAACTTCTCTGTGGGGGAGAGGCAGCTGCTCTGCATTGCCAGGGCTGTG	4226
Db	 4299	GAARACGGTGGAAACTTCTCTGTGGGGGAGAGGCAGCTGCTCTGCATTGCCAGGGCTGTG	4358
Qy	 4227	CTTCGCAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCTCCATTGACATGGAGACA	4286
Db	 4359	CTTCGCAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCTCCATTGACATGGAGACA	4418
Qy	 4287	GACACCTGATCCAGCGCACAAATCCCGTGAAGCCTTCCAGGCTGCACCCGTGCTCGTCATT	4346
Db	 4419	GACACCTGATCCAGCGCACAAATCCCGTGAAGCCTTCCAGGCTGCACCCGTGCTCGTCATT	4478
Qy	 4347	GCCACCGTGTCACCACTGTGTGAACCTGTGACCCACATCCTGGTTATGGGCAATGGGAAG	4406
Db	 4479	GCCACCGTGTCACCACTGTGTGAACCTGTGACCCACATCCTGGTTATGGGCAATGGGAAG	4538
Qy	 4407	GTGTAGAAATTTGATCGCGCGGAGGTACTGCGGAAGAAGCCTGGGTCAATTGTTCCGACCC	4466
Db	 4539	GTGTAGAAATTTGATCGCGCGGAGGTACTGCGGAAGAAGCCTGGGTCAATTGTTCCGACCC	4598
Qy	 4467	CTCATGGCCACAGCCACTTCTTCACTGAGATAAGGAGATGTGGAGACTTCATGGAGGCTG	4526
Db	 4599	CTCATGGCCACAGCCACTTCTTCACTGAGATAAGGAGATGTGGAGACTTCATGGAGGCTG	4658
Qy	 4527	GCAGCTGAGCTCAGAGGTTCCACACAGGTGCAGCTTCGAGGCCCCACAGTCTGCGACCTTCT	4586
Db	 4659	GCAGCTGAGCTCAGAGGTTCCACACAGGTGCAGCTTCGAGGCCCCACAGTCTGCGACCTTCT	4718
Qy	 4587	TGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAATGTAGGGGGGTGGGGAATTGC	4646
Db	 4719	TGTTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAATGTAGGGGGGTGGGGAATTGC	4778
Qy	 4647	TGGATGGAAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCCAGAACC	4706
Db	 4779	TGGATGGAAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCCAGAACC	4838
Qy	 4707	ATCTAAGACATGGGATTACAGTGATCATGTGGTTCTCCCTTTTAACTTACATGCTGAATAAT	4766
Db	 4839	ATCTAAGACATGGGATTACAGTGATCATGTGGTTCTCCCTTTTAACTTACATGCTGAATAAT	4898
Qy	 4767	TTTATAATAAGGTAAAAGCTTATAGTTTCTGATCTGTGTAGAAGTGTGCAAAATGCTG	4826
Db	 4899	TTTATAATAAGGTAAAAGCTTATAGTTTCTGATCTGTGTAGAAGTGTGCAAAATGCTG	4958
Qy	 4827	TACTGACTTTGTAAAAATATAAACTAAGGAAAACTC	4862

Db	4959	TACTGACTTTTGTAATAATATAAACTAAGGAAAACTC	4994
RESULT 8			
AX411822			
LOCUS	AX411822	4074 bp	DNA linear PAT 14-JUN-2002
DEFINITION	Sequence 2 from Patent WO0224742.		
ACCESSION	AX411822		
VERSION	AX411822.1	GI:21444310	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Krasnow, R.E. and Baughn, M.R.		
TITLE	Atp-binding cassette protein		
JOURNAL	Patent: WO 0224742-A 2 28-MAR-2002;		
	Incyte Genomics, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..4074		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/note="Incyte ID No: 7481813CB1"		
BASE COUNT	961 a 1050 c 1080 g 983 t		
ORIGIN			
	Query Match 81.5%; Score 3964.2; DB 6; Length 4074;		
	Best Local Similarity 99.8%; Pred No. 0;		
	Matches 3969; Conservative 0; Mismatches 8; Indels 0; Gaps 0;		
Qy	601	AGCCCTGGACAAATGCTGGCCGTCTCTCTACCTCACCGTGTCATGGCTCACCCCGCTCA	660
Db	98	ATCCCTGGACAAATGCTGGCCGTCTCTCTACCTCACCGTGTCATGGCTCACCCCGCTCA	157
Qy	561	TGATCCAAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTGTCAGTCCATG	720
Db	158	TGATCCAAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTGTCAGTCCATG	217
Qy	721	ATGCCTCAGACAAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAAGTCTCAAGGC	780
Db	218	ATGCCTCAGACAAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAAGTCTCAAGGC	277
Qy	781	GAGGATTGAAAAAGCTTCAGTGTCTTGTGTGATGCTGAGGTTCCAGAGAACAAAGTTGA	840
Db	278	GAGGATTGAAAAAGCTTCAGTGTCTTGTGTGATGCTGAGGTTCCAGAGAACAAAGTTGA	337
Qy	841	TTTTCGATGCACCTTCCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGGGCCAATATGA	900
Db	338	TTTTCGATGCACCTTCCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGGGCCAATATGA	397
Qy	901	TTATACCAAAGATCCTGGAATATTCAGAAGAGCAGTTGGGAATGTTGTCCATGGAGTGG	960
Db	398	TTATACCAAAGATCCTGGAATATTCAGAAGAGCAGTTGGGAATGTTGTCCATGGAGTGG	457
Qy	961	GACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTCCTCCAGTT	1020
Db	458	GACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTCCTCCAGTT	517
Qy	1021	GGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCCTCTTGGCTTTG	1080
Db	518	GGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCCTCTTGGCTTTG	577
Qy	1081	AGAGCTCATCCAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCCATCAGCTTCT	1140
Db	578	AGAGCTCATCCAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGTAATCAGCTTCT	637
Qy	1141	TCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGACACCCCTAGTACTGATCA	1200
Db	638	TCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGACACCCCTAGTACTGATCA	697
Qy	1201	CCTCGGCATCGCTGGTCATCTGCACCATTTCTTCCCTACTTCATATTATGGATACACTGCAT	1260

Db	698	CCTGGCCATCGTGGTCATCTGCAGCAATTTCTCCTACTTTCATTTATTGGATACACTGCAT	757
QY	1261	TTATTGCCATCTPATGCTATCTCCTGGTTTTCOCACCTGGCGGTATTTCATGACAAGAATGG	1320
Db	758	TTATTGCCATCTPATGCTATCTCCTGGTTTTCOCACCTGGCGGTATTTCATGACAAGAATGG	817
QY	1321	CTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCAGCGCATCCGTGTGACCAGTG	1380
Db	818	CTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCAGCGCATCCGTGTGACCAGTG	877
QY	1381	AAGTTCTCACTTGCATTAAAGCTGATTAAATGTACACATGGGAGAAACCATTTGCAAAAA	1440
Db	878	AAGTTCTCACTTGCATTAAAGCTGATTAAATGTACACATGGGAGAAACCATTTGCAAAAA	937
QY	1441	TCATTGAAGACCTTAAGAAGGAAGAAAGGAAACTATTGGAGAAGTGGGGCTTGTCCAGA	1500
Db	938	TCATTGAAGACCTTAAGAAGGAAGAAAGGAAACTATTGGAGAAGTGGGGCTTGTCCAGA	997
QY	1501	GCCTGACAAGTATAACCTTGTTCATCATCCOCACAGTGGCCACAGCGTCTGGGTTCTCA	1560
Db	998	GCCTGACAAGTATAACCTTGTTCATCATCCOCACAGTGGCCACAGCGTCTGGGTTCTCA	1057
QY	1561	TCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATGCTGGCCTCCT	1620
Db	1058	TCCACACATCCTTAAAGCTGAAACTCGCAGCGTCAATGGCCTTCAGCATGCTGGCCTCCT	1117
QY	1621	TGAATCTCCTTCGSGCTGTCAGTGTTCTTTTGTGCCTATTGCAGTCAAAGGTCTCACGAATT	1680
Db	1118	TGAATCTCCTTCGSGCTGTCAGTGTTCTTTTGTGCCTATTGCAGTCAAAGGTCCCACGAATT	1177
QY	1681	CCAAAGTCTGCAGTGATGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCTGTTTCTATG	1740
Db	1178	CTAAGTCTGCAGTGATGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCTGTTTCTATG	1237
QY	1741	TCCAGACATTACAAGACCCCCAGCAAAGCTCTGGTCTTTTGAGGAGGCCACCTTGTCAATGGC	1800
Db	1238	TCCAGACATTACAAGACCCCCAGCAAAGCTCTGGTCTTTTGAGGAGGCCACCTTGTCAATGGC	1297
QY	1801	AACAGACCTGTCCGGGGATCGTCAATGGGCACTGGAGCTGGAGAGGAACGGGCATGCTT	1860
Db	1298	AACAGACCTGTCCGGGGATCGTCAATGGGCACTGGAGCTGGAGAGGAACGGGCATGCTT	1357
QY	1861	CTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGAAGCAAGCCTGG	1920
Db	1358	CTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGAAGCAAGCCTGG	1417
QY	1921	GCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATGTAGGGGTCTGCG	1980
Db	1418	GCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATGTAGGGGTCTGCG	1477
QY	1981	GCAACACGGGAGTGGTAAGAGCAGCCCTGTGTGAGCCATCCTGGAGGAGATGCATTGC	2040
Db	1478	GCAACACGGGAGTGGTAAGAGCAGCCCTGTGTGAGCCATCCTGGAGGAGATGCATTGC	1537
QY	2041	TCGAGGGCTCGTGGGGGTGCAGGGAAGCCCTGGCCCTATGTCCCCCAGCAGGCCCTGGATCG	2100
Db	1538	TCGAGGGCTCGTGGGGGTGCAGGGAAGCCCTGGCCCTATGTCCCCCAGCAGGCCCTGGATCG	1597
QY	2101	TCAGCGGGAACATCAGGGAGAACATCCTCATGGGAGGCGCATATGACAAGGCCCGATACC	2160
Db	1598	TCAGCGGGAACATCAGGGAGAACATCCTCATGGGAGGCGCATATGACAAGGCCCGATACC	1657
QY	2161	TCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACCTTCTGCCCTTTGGAGACA	2220
Db	1658	TCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACCTTCTGCCCTTTGGAGACA	1717
QY	2221	TGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGCAGAAACAGAGGATCAGCC	2280
Db	1718	TGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGCAGAAACAGAGGATCAGCC	1777
QY	2281	TGGCCCGCGCGCTATTCGACCGTCAAGATCTACCTGCTGGACGACCCCTGCTGCTG	2340
Db	1778	TGGCCCGCGCGCTATTCGACCGTCAAGATCTACCTGCTGGACGACCCCTGCTGCTG	1837

QY	2341	TGACGCCCCACGTGGGAAGCACACATTTTTCAGGAGTGCATTTAAGAAGACACTCAGGGGGA	2400
Db	1838	TGACGCCCCACGTGGGAAGCACACATTTTTCAGGAGTGCATTTAAGAAGACACTCAGGGGGA	1897
QY	2401	AGACGGTCTGCTGTTGACCCACCAGCTGCAGTACTTAGAATTTTGTGGCAGATCATTT	2460
Db	1898	AGACGGTCTGCTGTTGACCCACCAGCTGCAGTACTTAGAATTTTGTGGCAGATCATTT	1957
QY	2461	TGTTGAAAAATGGAAAAATCTGTGAAAAATGAAACTCACAGTGAGTTAATGCAGAAAAAGG	2520
Db	1958	TGTTGAAAAATGGAAAAATCTGTGAAAAATGAAACTCACAGTGAGTTAATGCAGAAAAAGG	2017
QY	2521	GGAATATGCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGG	2580
Db	2018	GGAATATGCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGG	2077
QY	2581	ACACAGCAAAGATAGCAGAGAAAGCCAAAGGTAGAAAGTCAGGCTCTGGCCACCTCCCTGG	2640
Db	2078	ACACAGCAAAGATAGCAGAGAAAGCCAAAGGTAGAAAGTCAGGCTCTGGCCACCTCCCTGG	2137
QY	2641	AAGAGTCTCTCAACGGGAAATGCTGTGCCGGAGCATACACTACATCCAGGAGGAGAGATGG	2700
Db	2138	AAGAGTCTCTCAACGGGAAATGCTGTGCCGGAGCATACACTACATCCAGGAGGAGAGATGG	2197
QY	2701	AAGAAGGCTCCTTGAGTTGGAGGGTCTACCACTACATCCAGGAGGAGGAGGTTACA	2760
Db	2198	AAGAAGGCTCCTTGAGTTGGAGGGTCTACCACTACATCCAGGAGGAGGTTACA	2257
QY	2761	TGGTCTCTTGCATAAATTTCTTCTCGTGGTGCTGATCGTCTTCTTAACGATCTTCAGCT	2820
Db	2258	TGGTCTCTTGCATAAATTTCTTCTCGTGGTGCTGATCGTCTTCTTAACGATCTTCAGCT	2317
QY	2821	TCTGGTGGCTGAGTACTGGTTTGAGCAGGSGCTCGGGACCAATAGCAGCCGAGAGAGCA	2880
Db	2318	TCTGGTGGCTGAGTACTGGTTTGAGCAGGSGCTCGGGACCAATAGCAGCCGAGAGAGCA	2377
QY	2881	ATGGAACCATGSCAGACCTGGGCAACATTGCAGACAATCCTCAACTGTCTCTACCAAGC	2940
Db	2378	ATGGAACCATGSCAGACCTGGGCAACATTGCAGACAATCCTCAACTGTCTCTACCAAGC	2437
QY	2941	TGGTGTAACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGGTCTGCTCCTCAGGGATTT	3000
Db	2438	TGGTGTAACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGGTCTGCTCCTCAGGGATTT	2497
QY	3001	TCACCAAAGTCACGAGGAAGGCATCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTT	3060
Db	2498	TCACCAAAGTCACGAGGAAGGCATCCACGGCCCTGCACAACAAGCTCTTCAACAAGGTTT	2557
QY	3061	TCCGCTGCCCATGAGTTTCTTTGACACCAATCCCAATAGGCCGGCTTTTGAACCTGCTTCG	3120
Db	2558	TCCGCTGCCCATGAGTTTCTTTGACACCAATCCCAATAGGCCGGCTTTTGAACCTGCTTCG	2617
QY	3121	CAGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAGCAGTTCCTGGTCC	3180
Db	2618	CAGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAGCAGTTCCTGGTCC	2677
QY	3181	TGTCCCTTAATGGTGATCGCCGTCCCTGTTGATTGTGCTGCTGCCATATATCCTGT	3240
Db	2678	TGTCCCTTAATGGTGATCGCCGTCCCTGTTGATTGTGCTGCTGCCATATATCCTGT	2737
QY	3241	TAATGGGAGCCATAATCATGGTTATTGCTTCATTTATTATATGATGTTCAAGAAGGCCA	3300
Db	2738	TAATGGGAGCCATAATCATGGTTATTGCTTCATTTATTATATGATGTTCAAGAAGGCCA	2797
QY	3301	TCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTATGATGTTCAAGAAGGCCA	3360
Db	2798	TCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTATGATGTTCAAGAAGGCCA	2857
QY	3361	ATTCTCTGCAAGGCCCTGAGCTCCATCCATCCTATGGAAAAACTGAAGACTTCATCAGCC	3420
Db	2858	ATTCTCTGCAAGGCCCTGAGCTCCATCCATCCTATGGAAAAACTGAAGACTTCATCAGCC	2917

Qy	3421	AGTTAAAGAGGCTGACTGATGCGCAGAATAACTACCTGCTGTTGTTTCTATCTTCCACAC	3480
Db	2918	AGTTTAAAGAGGCTGACTGATGCGCAGAATAACTACCTGCTGTTGTTTCTATCTTCCACAC	2977
Qy	3481	GATGGATGGCATTTGAGGCTGGAGATCATGACCAACCTTTGACCTTGGCTGTTGCCCTGT	3540
Db	2978	GATGGATGGCATTTGAGGCTGGAGATCATGACCAACCTTTGACCTTGGCTGTTGCCCTGT	3037
Qy	3541	TCGTGGCTTTTGGCATTTCCCTCCACCCCTACTCCTTTAAAGTCATGGCTGTCAACATCG	3600
Db	3038	TCGTGGCTTTTGGCATTTCCCTCCACCCCTACTCCTTTAAAGTCATGGCTGTCAACATCG	3097
Qy	3601	TGCTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCCCGATTGGCTTGGAGACAGAGGCAC	3660
Db	3098	TGCTGCAGCTGGCGTCCAGCTTCCAGGCCACTGCCCGATTGGCTTGGAGACAGAGGCAC	3157
Qy	3661	AGTTCACGGCTGTAGAGAGGATFACTGTCAGTACATGAAGATGTGTCTCGGAAGCTCCTT	3720
Db	3158	AGTTCACGGCTGTAGAGAGGATFACTGTCAGTACATGAAGATGTGTCTCGGAAGCTCCTT	3217
Qy	3721	TACACATGGAAGGCACAAGTTGTCCTCCAGGGGTGGCCACAGCATGGGAAATCATATTTT	3780
Db	3218	TACACATGGAAGGCACAAGTTGTCCTCCAGGGGTGGCCACAGCATGGGAAATCATATTTT	3277
Qy	3781	AGGATTATCACATGAAATACAGAGACAACACACCCACCGTGCTTCACGGCATCAACCTGA	3840
Db	3278	AGGATTATCACATGAAATACAGAGACAACACACCCACCGTGCTTCACGGCATCAACCTGA	3337
Qy	3841	CCATCCGCGGCCACGAAGTGTGGGCATCGTGGGAAGACGGGCTCTGGGAAGTCTCCT	3900
Db	3338	CCATCCGCGGCCACGAAGTGTGGGCATCGTGGGAAGACGGGCTCTGGGAAGTCTCCT	3397
Qy	3901	TGGCATGGCTCTCTTCCGCTGGTGGAGCCCATGGCAGCCCGGATTTCTCATTTAGCGCG	3960
Db	3398	TGGCATGGCTCTCTTCCGCTGGTGGAGCCCATGGCAGCCCGGATTTCTCATTTAGCGCG	3457
Qy	3961	TGGACATTTGCAGCATCGGCTGGAGGACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAG	4020
Db	3458	TGGACATTTGCAGCATCGGCTGGAGGACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAG	3517
Qy	4021	ATCCAGTGTGCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTTGACCCGTCACTG	4080
Db	3518	ATCCAGTGTGCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTTGACCCGTCACTG	3577
Qy	4081	ACCAGCAGATCTGGGATGCCCTTGGAGAGGACATTCCTGACCAGGCCATCTCAAAGTTCC	4140
Db	3578	ACCAGCAGATCTGGGATGCCCTTGGAGAGGACATTCCTGACCAGGCCATCTCAAAGTTCC	3637
Qy	4141	CCAAAAGCTGCATACAGATGTGGTGGAAAACGGTGGAAACTTCTCTGTGGGGAGAGGC	4200
Db	3638	CCAAAAGCTGCATACAGATGTGGTGGAAAACGGTGGAAACTTCTCTGTGGGGAGAGGC	3697
Qy	4201	AGCTGCTCTGCATTTGCCAGGCTGTGCTTCGCAACTCCAAGATCATCCTTATCGATGAAG	4260
Db	3698	AGCTGCTCTGCATTTGCCAGGCTGTGCTTCGCAACTCCAAGATCATCCTTATCGATGAAG	3757
Qy	4261	CCACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATCCGTGAAGCCT	4320
Db	3758	CCACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAAATCCGTGAAGCCT	3817
Qy	4321	TCCAGGGCTGCACCGTGCTCGTCATTTGCCACCCGTGCACCACCTGTGCTGAACCTGACC	4380
Db	3818	TCCAGGGCTGCACCGTGCTCGTCATTTGCCACCCGTGCACCACCTGTGCTGAACCTGACC	3877
Qy	4381	ACATCCTGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCGGAGGTACTGCGGA	4440
Db	3878	GCATCCTGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCGGAGGTACTGCGGA	3937
Qy	4441	AGAAGCCTGGGTCAATTGTCGACGCCCTCATGGCCACAGCCACTTCTTCACTGAGATAAG	4500
Db	3938	AGAAGCCTGGGTCAATTGTCGACGCCCTCATGGCCACAGCCACTTCTTCACTGAGATAAG	3997
Qy	4501	GAGATGTGGAGACTTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTCCACACAGGTCAGCT	4560

Db	3998	GAGATGTGGAGACTTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTTCACACAGGTGCAGCT	4057
QY	4561	TCGAGGCCACAGTCTG	4577
Db	4058	TCGAGGCCACAGTCTG	4074
RESULT 9			
AX468402			
LOCUS	AX468402	4083 bp	DNA linear PAT 15-JUL-2002
DEFINITION	Sequence 3 from Patent WO0216589.		
ACCESSION	AX468402		
VERSION	AX468402.1 GI:21901255		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Curtis,R.A.		
TITLE	Human ABC transporter family member and uses thereof		
JOURNAL	Patent: WO 0216589-A 3 28-FEB-2002;		
Millennium Pharmaceuticals, Inc. (US)			
FEATURES	Location/Qualifiers		
source	1..4083		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	954 a	1061 c	1078 g 990 t
ORIGIN			
Query Match 78.9%; Score 3836.4; DB 6; Length 4083;			
Best Local Similarity 96.1%; Pred. No. 0;			
Matches 4023; Conservative 0; Mismatches 21; Indels 144; Gaps 3;			
QY	351	ATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	410
Db	1	ATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	60
QY	411	ATCGACATAGGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT	470
Db	61	ATCGACATAGGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT	120
QY	471	GGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGAGGGCAGCTGTCCCACCG	530
Db	121	GGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGAGGGCAGCTGTCCCACCG	180
QY	531	TGGGGGAAGTATGATGCTGCCTTGAGAAACCATGATTCCCTTCCGTCCTCCCAAGCCGAGGTTT	590
Db	181	TGGGGGAAGTATGATGCTGCCTTGAGAAACCATGATTCCCTTCCGTCCTCCCAAGCCGAGGTTT	240
QY	591	CCTGCCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTTCCTACCTCACCGTGTATGGCTC	650
Db	241	CCTGCCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTTCCTACCTCACCGTGTATGGCTC	300
QY	651	ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAAACCATCCCTCCACTG	710
Db	301	ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAAACCATCCCTCCACTG	360
QY	711	TCAGTCCATGATGCCTCAGACAAAATGTCCAAAGGCTTCACCGCTTTGGGAAGAAGAA	770
Db	361	TCAGTCCATGATGCCTCAGACAAAATGTCCAAAGGCTTCACCGCTTTGGGAAGAAGAA	420
QY	771	GTCTCAAGGCGAGGGATTGAAAAGCTTCAGTGCTTCTGGTGTATGCTGAGGTTCCAGAGA	830
Db	421	GTCTCAAGGCGAGGGATTGAAAAGCTTCAGTGCTTCTGGTGTATGCTGAGGTTCCAGAGA	480
QY	831	ACAAGGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGGG	890
Db	481	ACAAGGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGGG	540
QY	891	CCAATATTGATTATACCAAGATCCTGGAATATTCAAGAGACCAGTTGGGGAATGTTGTC	950

Db 541 CCAATATTGATTATACCAAAGATCCTGGAAATATTACAGAAGAGCAGTGGGGAATGTTGTC 600

QY 951 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGGAAGTCTCTGAGTTTC 1010

Db 601 CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGGAAGTCTCTGAGTTTC 660

QY 1011 TCCTCCAGTTGGATCATCAACCAACGACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCC 1070

Db 661 TCCTCCAGTTGGATCATCAACCAACGACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCC 720

QY 1071 TTTGGCCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATACACATCACCTCAGGAG--- 1125

Db 721 TTTGGCCTTTGAGAAGCTCATCCAATTTAAGTCTGTAATACACATCACCTCAGGAGGGA 780

QY 1126 -----AGGCCATCAGCTTCTTACCGGTGAT 1151

Db 781 GGTGACATCTGTGCCCATCAACTTGCTGTCTTGACGGCCATCAGCTTCTTACCGGTGAT 840

QY 1152 GTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTAGTACTGATCACCTGCGCATCG 1211

Db 841 GTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTAGTACTGATCACCTGCGCATCG 900

QY 1212 CTGGTCACTGCAGCATTTCTTCCCTACTTCAATTATTGGATACACTGCATTTATTGCCATC 1271

Db 901 CTGGTCACTGCAGCATTTCTTCCCTACTTCAATTATTGGATACACTGCATTTATTGCCATC 960

QY 1272 TTATGCTATCTCCTGGTTTTTCCCACCTGGCGGTATTTCATGACAAAGAAATGGCTGTGAAGGCT 1331

Db 961 TTATGCTATCTCCTGGTTTTTCCCACCTGGCGGTATTTCATGACAAAGAAATGGCTGTGAAGGCT 1020

QY 1332 CAGCATCACACATCTGAGGTGACGGACCGCATCCGTTGTGACCAGTGAAGTTCTCACT 1391

Db 1021 CAGCATCACACATCTGAGGTGACGGACCGCATCCGTTGTGACCAGTGAAGTTCTCACT 1080

QY 1392 TGCATTAACTGATTAAAAATGTACACATGGGAGAAACCATTTTGCANAAATCATTTGAAGAC 1451

Db 1081 TGCATTAACTGATTAAAAATGTACACATGGGAGAAACCATTTTGCANAAATCATTTGAAG-- 1138

QY 1452 CTAAGAGGAAGGAAGAACTATTGGAGAAAGTGGGGCTTGTCCAGAGCCTGACAAGT 1511

Db 1139 -----GTATGGAAAGT 1149

QY 1512 ATAACCTTGTTCATCATCCCCACAGTGGCCACACAGCGTCTGGGTCTCTCATCCACACATCC 1571

Db 1150 CTGACTTCTGCTCCAAACCTGGTGATGGC----- 1179

QY 1572 TTAAAGCTGAACACTCACAGCGTCAATGGCCCTTCAGCATGCTGGCTCCTTGAATCTCCTT 1631

Db 1180 -----ATGGCCTTCAGCATGCTGGCCTCCTTGAATCTCCTT 1215

QY 1632 CGGCTGTCACTGTTCTTTGTGCCATTATGCAGTCAAAGGTCTCACGAATTTCCAAAGTCTGCA 1691

Db 1216 CGGCTGTCACTGTTCTTTGTGCCATTATGCAGTCAAAGGTCTCACGAATTTCCAAAGTCTGCA 1275

QY 1692 GTGATGAGGTTCAAGAAAGTTTTTTCCTCCAGGAGAGCCCTGTTTCTATGTCCAGACATTA 1751

Db 1276 GTGATGAGGTTCAAGAAAGTTTTTTCCTCCAGGAGAGCCCTGTTTCTATGTCCAGACATTA 1335

QY 1752 CAAGACCCAGCAAAGCTCTGGTCTTTGAGGAGGCCACCTTGTATGGCAACAGACCTGT 1811

Db 1336 CAAGACCCAGCAAAGCTCTGGTCTTTGAGGAGGCCACCTTGTATGGCAACAGACCTGT 1395

QY 1812 CCCGGATCGTCAATGGGCACTTGAGCTGGAGAGGAACAGCCCTGGGCCAGAGTTG 1871

Db 1396 CCCGGATCGTCAATGGGCACTTGAGCTGGAGAGGAACGGGATGCTTCTGAGGGGATG 1455

QY 1872 ACCAGGCTAGAGATGCCCTCGGGCCAGAGGAAGGAAGGAACAGCCCTGGGCCAGAGTTG 1931

Db 1456 ACCAGGCTAGAGATGCCCTCGGGCCAGAGGAAGGAAGGAACAGCCCTGGGCCAGAGTTG 1515

QY 1932 CACAAGATCAACCTGGTGGTGTCCAAGGGGATGATGTTAGGGTCTGCGGCAACACGGGG 1991

Db 1516 CACAAGATCAACCTGGTGGTGTCCAAGGGGATGATGTTAGGGTCTGCGGCAACACGGGG 1575

QY 1992 AGTGTAAAGACAGACCCCTGTTGTACCCCATCCTGGAGGAGATGCACTTGCTCGAGGGCTCG 2051

Db 1576 AGTGTAAAGACAGACCCCTGTTGTACCCCATCCTGGAGGAGATGCACTTGCTCGAGGGCTCG 1635

QY 2052 GTGGGGTGCAGGGAAGCCTGGCCTATGTCCCCCAGCAGGCCCTGGATCGTCAGCGGAAC 2111

Db 1636 GTGGGGTGCAGGGAAGCCTGGCCTATGTCCCCCAGCAGGCCCTGGATCGTCAGCGGAAC 1695

QY 2112 ATCAGGAGAAACATCCTCATGGAGGCGCATATGACAAGGCCGATACCTCCAGGTGCTC 2171

Db 1696 ATCAGGAGAAACATCCTCATGGAGGCGCATATGACAAGGCCGATACCTCCAGGTGCTC 1755

QY 2172 CACTGCTGCTCCCTGAATCGGGACCTGGAACCTTCTGCCCTTTGGAGACATGACAGAGATT 2231

Db 1756 CACTGCTGCTCCCTGAATCGGGACCTGGAACCTTCTGCCCTTTGGAGACATGACAGAGATT 1815

QY 2232 GGAGAGCGGGCCTCAACCTCTCTGGGGGCGAGAAACAGAGGATCAGCCTGGCCCCGCC 2291

Db 1816 GGAGAGCGGGCCTCAACCTCTCTGGGGGCGAGAAACAGAGGATCAGCCTGGCCCCGCC 1875

QY 2292 GTCTATTCCGACCGTCAAGTCTACCTGTGGACGACCCCTGTCTGTGTGGACGCCAC 2351

Db 1876 GTCTATTCCGACCGTCAAGTCTACCTGTGGACGACCCCTGTCTGTGTGGACGCCAC 1935

QY 2352 GTGGGAAGCACATTTTGTAGGAGTGCATTAAGAAGACACTCAGGGGGAAGACGGTCGTC 2411

Db 1936 GTGGGAAGCACATTTTGTAGGAGTGCATTAAGAAGACACTCAGGGGGAAGACGGTCGTC 1995

QY 2412 CTGTGACCCACCACTGTCAGTGCATTTAGAAATTTTGTGGCCAGATCATTTTGTGAAAAT 2471

Db 1996 CTGTGACCCACCACTGTCAGTGCATTTAGAAATTTTGTGGCCAGATCATTTTGTGAAAAT 2055

QY 2472 GGGAAATCTGTGAAATGGAATCAGAGTGCATTAATGCAGAAAAAGGGAAATATGCC 2531

Db 2056 GGGAAATCTGTGAAATGGAATCAGAGTGCATTAATGCAGAAAAAGGGAAATATGCC 2115

QY 2532 CAACCTATCCAGAAAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGGACACAGCAAAG 2591

Db 2116 CAACCTATCCAGAAAGATGCACAAGGAAGCCACTTCGGACATGTTGCAGGACACAGCAAAG 2175

QY 2592 ATAGCAGAGAAGCCCAAAGGTAGAAAGTCAGGCTCTGGCCACCTCCCTGGAAGAGTCTCTC 2651

Db 2176 ATAGCAGAGAAGCCCAAAGGTAGAAAGTCAGGCTCTGGCCACCTCCCTGGAAGAGTCTCTC 2235

QY 2652 AACGGAAATGCTGTGCGGGAGCATCAGCTCACACAGGAGGAGAGATGGAAGAAGGCTCC 2711

Db 2236 AACGGAAATGCTGTGCGGGAGCATCAGCTCACACAGGAGGAGAGATGGAAGAAGGCTCC 2295

QY 2712 TTGAGTTGAGGGTCTACCACTACATCCAGGCAGCTGGAGGTACATGGTCTCTTGC 2771

Db 2296 TTGAGTTGAGGGTCTACCACTACATCCAGGCAGCTGGAGGTACATGGTCTCTTGC 2355

QY 2772 ATAAATTTCTTCTTCGTGGTGCATCGTCTTCTTAACGATCTTCAGCTTCTGGTGGCTG 2831

Db 2356 ATAAATTTCTTCTTGTGGTGCATCGTCTTCTTAACGATCTTCAGCTTCTGGTGGCTG 2415

QY 2832 AGCTACTGTTGGAGCAGGGCTCGGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATG 2891

Db 2416 AGCTACTGTTGGAGCAGGGCTCGGGGACCAATAGCAGCCGAGAGAGCAATGGAACCATG 2475

QY 2892 GCAGACCTGGGCAACATTGCAGACAAATCCTCAACTGTCTCTTACCAGCTGGTACGGG 2951

Db 2476 GCAGACCTGGGCAACATTGCAGACAAATCCTCAACTGTCTCTTACCAGCTGGTACGGG 2535

QY 2952 CTCACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCCTCAGGGATTTTCAACAAAGTC 3011

Db 2536 CTCACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCCTCAGGGATTTTCAACAAAGTC 2595

QY 3012 ACGAGAAGGCATCCACGGCCCTGCACAAACAAAGCTCTTCAACAAAGTTTTCGCTGCC 3071

Db 2596 ACGAGAAGGCATCCACGGCCCTGCACAAACAAAGCTCTTCAACAAAGTTTTCGCTGCC 2655

QY 1608 ATGCTGGCCTCCTTGAAATCTCCTTCGGCTGTCAAGTGTCTTTGTCGCTATTGCAGTCAA 1667
Db 301 ATGCTGGCCTCCTTGAAATCTCCTTCGGCTGTCAAGTGTCTTTGTCGCTATTGCAGTCAA 360
QY 1668 GGTCTACGAATCCAAAGTCTGCAGTGATGAGGTTCAAGAAGTTTTTCTCCAGGAGAGC 1727
Db 361 GGTCACGAATCCAAAGTCTGCAGTGATGAGGTTCAAGAAGTTTTTCTCCAGGAGAGC 420
QY 1728 CCTGTTTCTATGTCCAGACATTACAAGACCCCCAGCAAAGCTCTGGTCTTTGAGGAGGCC 1787
Db 421 CCTGTTTCTATGTCCAGACATTACAAGACCCCCAGCAAAGCTCTGGTCTTTGAGGAGGCC 480
QY 1788 ACCTTGTCATGGCAACAGACCTGTCCCGGGATCGTCAATGGGGCACTGGAGCTGGAGAGG 1847
Db 481 ACCTTGTCATGGCAACAGACCTGTCCCGGGATCGTCAATGGGGCACTGGAGCTGGAGAGG 540
QY 1848 AACGGGCATGCTTCTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGAA 1907
Db 541 AACGGGCATGCTTCTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGAA 600
QY 1908 GGAACAGCTGGGCCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCTCAAGGGGATGATG 1967
Db 601 GGAACAGCTGGGCCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCTCAAGGGGATGATG 660
QY 1968 TTAGGGGTCTCGGGCAACACGGGGAGTGGTAAGAGCAGCCCTGTTGTCAGCCATCCTGGAG 2027
Db 661 TTAGGGGTCTCGGGCAACACGGGGAGTGGTAAGAGCAGCCCTGTTGTCAGCCATCCTGGAG 720
QY 2028 GAGATGCACTGTCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCCTGGCCTATGTCCCCCAG 2087
Db 721 GAGATGCACTGTCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCCTGGCCTATGTCCCCCAG 780
QY 2088 CAGGCCCTGGATCGTCAGCGGGGAACATCAGGGAGAACATCCTCATGGGAGCGCATATGAC 2147
Db 781 CAGGCCCTGGATCGTCAGCGGGGAACATCAGGGAGAACATCCTCATGGGAGCGCATATGAC 840
QY 2148 AAGGCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACCTCTG 2207
Db 841 AAGGCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACCTCTG 900
QY 2208 CCCTTTGGAGACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGCGAGAA 2267
Db 901 CCCTTTGGAGACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGCGAGAA 960
QY 2268 CAGAGGATCAGCCTGGCCCCGGCCGCTCTATTCCGACCGTCAGATCTACCTGCTGGACGAC 2327
Db 961 CAGAGGATCAGCCTGGCCCCGGCCGCTCTATTCCGACCGTCAGATCTACCTGCTGGACGAC 1020
QY 2328 CCCCTGCTGCTGTGGACGCCCCACGTGGGGGAAGCACATTTTGGAGAGTGCATTAAGAAG 2387
Db 1021 CCCCTGCTGCTGTGGACGCCCCACGTGGGGGAAGCACATTTTGGAGAGTGCATTAAGAAG 1080
QY 2388 ACACCTCAGGGGAAGACGGTCTGCTGGTGACCCACAGCTGCAGTACTTAGAATTTTGT 2447
Db 1081 ACACCTCAGGGGAAGACGGTCTGCTGGTGACCCACAGCTGCAGTACTTAGAATTTTGT 1140
QY 2448 GGCCAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAATGGAATCAGTGAGTTA 2507
Db 1141 GGCCAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAATGGAATCAGTGAGTTA 1200
QY 2508 ATGCAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCG 2567
Db 1201 ATGCAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCG 1260
QY 2568 GACATGTCAGGACACAGCAAAGATAGCAGAGAAGCCAAAGGTAGAAAAGTCAGGCTCTG 2627
Db 1261 GACATGTCAGGACACAGCAAAGATAGCAGAGAAGCCAAAGGTAGAAAAGTCAGGCTCTG 1320
QY 2628 GCCACCTCCCTGGAAGAGTCTCTCAACGGGAAATGCTGTGCCGGAGCATCAGCTCACACAG 2687
Db 1321 GCCACCTCCCTGGAAGAGTCTCTCAACGGGAAATGCTGTGCCGGAGCATCAGCTCACACAG 1380
QY 2688 GAGGAGGAGATGGAAGAAGGCTCCTTGAGTTGAGGGTCTACCACCACTACATCCAGGCA 2747

Db 1381 GAGGAGGAGATGGAAGAAGGCTCCTTGAGTTGGAGGGTCTACCACCCTACATCCAGGCA 1440
QY 2748 GCTGGAGGTTACATGGTCTCTTGCATAAATTTCTTCTTCGTGGTGCTGATCGTCTTTA 2807
Db 1441 GCTGGAGGTTACATGGTCTCTTGCATAAATTTCTTCTTCGTGGTGCTGATCGTCTTTA 1500
QY 2808 ACGATCTTACGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGC 2867
Db 1501 ACGATCTTACGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGC 1560
QY 2868 AGCCGAGAGACAAATGGAACCATGGCAGACCTGGCAACATTGCAGACAAATCCTCAACTG 2927
Db 1561 AGCCGAGAGACAAATGGAACCATGGCAGACCTGGCAACATTGCAGACAAATCCTCAACTG 1620
QY 2928 TCCTTCTACAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGC 2987
Db 1621 TCCTTCTACAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGC 1680
QY 2988 TCCTCAGGGATTTTCACCAAAGTTCACGAGGAAGGCATCCACGGCCCTGCACACAAGCTC 3047
Db 1681 TCCTCAGGGATTTTCACCAAAGTTCACGAGGAAGGCATCCACGGCCCTGCACACAAGCTC 1740
QY 3048 TTCAACAAGGTTTTCCCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGGCCGGCTT 3107
Db 1741 TTCAACAAGGTTTTCCCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGGCCGGCTT 1800
QY 3108 TTGAACCTGCTTCGAGGGGACTTTGGAACAGCTGGACCAAGCTCTTGCCCATCTTTTCAGAG 3167
Db 1801 TTGAACCTGCTTCGAGGGGACTTTGGAACAGCTGGACCAAGCTCTTGCCCATCTTTTCAGAG 1860
QY 3168 CAGTTCCTGGTCTGTCTTAATGGTGATCGCCGCTCCTGTTGATTGTGCTGTCT 3227
Db 1861 CAGTTCCTGGTCTGTCTTAATGGTGATCGCCGCTCCTGTTGATTGTGCTGTCT 1920
QY 3228 CCATATATCCTGTTAATGGGAGCCATAATCATCATGGTTATTGCTTCAATTATATATGATG 3287
Db 1921 CCATATATCCTGTTAATGGGAGCCATAATCATCATGGTTATTGCTTCAATTATATATGATG 1980
QY 3288 TTCAAGAAGGCCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGCTCTCCTTTATTC 3347
Db 1981 TTCAAGAAGGCCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGCTCTCCTTTATTC 2040
QY 3348 TCCCACATCCTCAATTCTCTGCAAGGCCCTGAGCTCCATCCATGTCTATGGAAAACGTAA 3407
Db 2041 TCCCACATCCTCAATTCTCTGCAAGGCCCTGAGCTCCATCCATGTCTATGGAAAACGTAA 2100
QY 3408 GACTTCATCAGCCAGTTTAAAGAGGCTGACTGATGGCAGAAATACTACCTGCTGTTT 3467
Db 2101 GACTTCATCAGCCAGTTTAAAGAGGCTGACTGATGGCAGAAATACTACCTGCTGTTT 2160
QY 3468 CTATCTTCCACACGATGGATGGCATTGAGGCTGGAGATCATGACCACCTTGTGACCTTG 3527
Db 2161 CTATCTTCCACACGATGGATGGCATTGAGGCTGGAGATCATGACCACCTTGTGACCTTG 2220
QY 3528 GCTGTTGCCCTGTTTCGTGGCTTTTGGCATTTCTCCACCCCTACTCCTTTAAAGTCATG 3587
Db 2221 GCTGTTGCCCTGTTTCGTGGCTTTTGGCATTTCTCCACCCCTACTCCTTTAAAGTCATG 2280
QY 3588 GCTGTCAACATCGTGTGCAGCTGGCGTCCAGCTTCCAGGGCCACTGCCCCGATTGGCTTG 3647
Db 2281 GCTGTCAACATCGTGTGCAGCTGGCGTCCAGCTTCCAGGGCCACTGCCCCGATTGGCTTG 2340
QY 3648 GAGACAGAGGCACAGTTTCACCGCTGTAGAGAGGATACTGCAGTACATGAAGATGTGTGC 3707
Db 2341 GAGACAGAGGCACAGTTTCACCGCTGTAGAGAGGATACTGCAGTACATGAAGATGTGTGC 2400
QY 3708 TCGGAAGCTCCTTTACACATGGAGGCACAAGTTGTCCCCAGGGGTGGCCACAGCATGGG 3767
Db 2401 TCGGAAGCTCCTTTACACATGGAGGCACAAGTTGTCCCCAGGGGTGGCCACAGCATGGG 2460
QY 3768 GAAATCATATTTTCAGGATTATCATATGAATACAGAGACAACACCCACCGTGTTCAC 3827

Db	2461	GAAATCATATTTCAGGATTATACACATGAAATACAGAGACAACACACCCACCGTGTTCAC	2520
QY	3828	GGCATCAACCTGACCATCCGCGGCCACGAAGTGGTGGGCATCGTGGGAAGGACGGGCTCT	3887
Db	2521	GGCATCAACCTGACCATCCGCGGCCACGAAGTGGTGGGCATCGTGGGAAGGACGGGCTCT	2580
QY	3888	GGGAAGTCTCCTTGGGCATGGCTCTCTTCCGCCTGGTGGAGCCCATGGCAGGCGGATT	3947
Db	2581	GGGAAGTCTCCTTGGGCATGGCTCTCTTCCGCCTGGTGGAGCCCATGGCAGGCGGATT	2640
QY	3948	CTCATTTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAGGACTTTCGGTCCAAGCTCTCA	4007
Db	2641	CTCATTTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAGGACTTTCGGTCCAAGCTCTCA	2700
QY	4008	GTGATCCCTCAAGATCCAGTGTCTCTCTCAGGAACCATCAGATTCACACCTAGATCCCTTT	4067
Db	2701	GTGATCCCTCAAGATCCAGTGTCTCTCTCAGGAACCATCAGATTCACACCTAGATCCCTTT	2760
QY	4068	GACCGTCACACTGACCCAGCAGATCTGGGATGCCCTTGGAGAGGACATTCCTTGACCAAGGCC	4127
Db	2761	GACCGTCACACTGACCCAGCAGATCTGGGATGCCCTTGGAGAGGACATTCCTTGACCAAGGCC	2820
QY	4128	ATCTCAAAGTTCCTCCAAAAAGCTGCATACAGATGTGGTGGAAAAACGGTGGAAACTTCTCT	4187
Db	2821	ATCTCAAAGTTCCTCCAAAAAGCTGCATACAGATGTGGTGGAAAAACGGTGGAAACTTCTCT	2880
QY	4188	GTGGGGGAGAGCGAGCTGTCTGTGCATTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATC	4247
Db	2881	GTGGGGGAGAGCGAGCTGTCTGTGCATTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATC	2940
QY	4248	CTTATCGATGAAGCCACAGCCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCAC	4307
Db	2941	CTTATCGATGAAGCCACAGCCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCAC	3000
QY	4308	ATCCGTGAAGCCTTCCAGGGCTGCACCGTGTCTCGTCATTGCCACCGTGTCAACCACTGTG	4367
Db	3001	ATCCGTGAAGCCTTCCAGGGCTGCACCGTGTCTCGTCATTGCCACCGTGTCAACCACTGTG	3060
QY	4368	CTGAAGTGTGACCCACATCCTTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCG	4427
Db	3061	CTGAAGTGTGACCCACATCCTTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCG	3120
QY	4428	GAGGTACTGCGGAAGAACCTGGGTCAATTGTTCCGAGCCCTCATGGCCACAGCCACTTCT	4487
Db	3121	GAGGTACTGCGGAAGAAGCCTGGGTCAATTGTTCCGAGCCCTCATGGCCACAGCCACTTCT	3180
QY	4488	TCACTGAGA	4496
Db	3181	TCACTGAGA	3189

RESULT 11
AX135190
LOCUS AX135190
DEFINITION Sequence 23 from Patent WO0132706.
ACCESSION AX135190
VERSION AX135190.1 GI:14271540
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3660)
AUTHORS Harras,M., Donoho,G., Turner,C.A., Nehls,M., Friedrich,G.,
Zambrowicz,B. and Sands,A.T.
TITLE Novel human transporter proteins and polynucleotides encoding the
same
JOURNAL Patent: WO 0132706-A 23 10-MAY-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..3660
/organism="Homo sapiens"
/db_xref="taxon:9606"

	BASE COUNT	862 a	952 c	955 g	891 t	
	ORIGIN					
	Query Match	64.9%;	Score 3156.2;	DB 6;	Length 3650;	
	Best Local Similarity	88.1%;	Pred. No. 0;			
	Matches 3657;	Conservative 0;	Mismatches 3;	Indels 489;	Gaps 1;	
QY	351	ATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	410			
Db	1	ATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	60			
QY	411	ATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT	470			
Db	61	ATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT	120			
QY	471	GGCCCTGGAGTCAGCAAGAGAGAAAATCCTGAGGCTCCAGGGAGGGCAGCTGTCCCACCG	530			
Db	121	GGCCCTGGAGTCAGCAAGAGAGAAAATCCTGAGGCTCCAGGGAGGGCAGCTGTCCCACCG	180			
QY	531	TGGGGGAAGTATGATGCTGCCTTGAGAAACCATGATTCCTCTCCGTCCCAAGCCGAGGTTT	590			
Db	181	TGGGGGAAGTATGATGCTGCCTTGAGAAACCATGATTCCTCTCCGTCCCAAGCCGAGGTTT	240			
QY	591	CCTGCCCCCCCAGCCCCCTGGACAATGCTGGCCCTGTCTCTTACCTCACCGTGTATGGCTC	650			
Db	241	CCTGCCCCCCCAGCCCCCTGGACAATGCTGGCCCTGTCTCTTACCTCACCGTGTATGGCTC	300			
QY	651	ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTG	710			
Db	301	ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTG	360			
QY	711	TCAGTCCATGATGCCTCAGACAAAAATGTCCAAGGCTTCACCGCCTTTGGGAAGAAGAA	770			
Db	361	TCAGTCCATGATGCCTCAGACAAAAATGTCCAAGGCTTCACCGCCTTTGGGAAGAAGAA	420			
QY	771	GTCTCAAGCGGAGGGATTGAAAAAGCTTCAGTGGTCTCTGGTATGCTGAGGTTCCAGAGA	830			
Db	421	GTCTCAAGCGGAGGGATTGAAAAAGCTTCAGTGGTCTCTGGTATGCTGAGGTTCCAGAGA	480			
QY	831	ACAAGGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGG	890			
Db	481	ACAAGGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGG	540			
QY	891	CCAATATTGATTATACCAAGATCCTGGAATATTCAGAAAGACAGATTGGGSAATGTTGTC	950			
Db	541	CCAATATTGATTATACCAAGATCCTGGAATATTCAGAAAGACAGATTGGGSAATGTTGTC	600			
QY	951	CATGGAGTGGGACTCTGCTTTGCCCTTTTTCCTCCGAATGTGTGAAGTCTCTGAGTTTC	1010			
Db	601	CATGGAGTGGGACTCTGCTTTGCCCTTTTTCCTCCGAATGTGTGAAGTCTCTGAGTTTC	660			
QY	1011	TCCTCCAGTTGGATCATCAACCACGCACAGCCATCAGGTTCCGAGCAGCTGTTCCTCC	1070			
Db	661	TCCTCCAGTTGGATCATCAACCACGCACAGCCATCAGGTTCCGAGCAGCTGTTCCTCC	720			
QY	1071	TTTGCCCTTTGAGAAAGCTCATCCAATTTAAGTCTGTAATACACATCACCTCAGGAGGCC	1130			
Db	721	TTTGCCCTTTGAGAAAGCTCATCCAATTTAAGTCTGTAATACACATCACCTCAGGAGGCC	780			
QY	1131	ATCAGCTTCTTCACCGGTGATGTAAACTACCTTGTGAGGGGTGTGCTATGACCCCTA	1190			
Db	781	ATCAGCTTCTTCACCGGTGATGTAAACTACCTTGTGAGGGGTGTGCTATGACCCCTA	840			
QY	1191	GTACTGATCACCTGCGCATCGCTGGTCACTCGCAGCATTTCTTCCCTACTTCATTATGGA	1250			
Db	841	GTACTGATCACCTGCGCATCGCTGGTCACTCGCAGCATTTCTTCCCTACTTCATTATGGA	900			
QY	1251	TACACTGCATTTATGCCCATCTTATGCTATCTCCTGGTTCCTCCACTGGCGGTATTCATG	1310			
Db	901	TACACTGCATTTATGCCCATCTTATGCTATCTCCTGGTTCCTCCACTGGAGGTATTCATG	960			
QY	1311	ACAAGAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT	1370			

Db	961	ACAAGATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGGACCAGCGCATCCGT	1020
Qy	1371	GTGACCACTGAAGTTCTCAGTTGCATTAAGCTGATTAAAAATGTACACATGGGAGAAACCA	1430
Db	1021	GTGACCACTGAAGTTCTCAGTTGCATTAAGCTGATTAAAAATGTACACATGGGAGAAACCA	1080
Qy	1431	TTTGCAAAAATCATTTGAAGACCTAAGAAGGAAGGAAAGGAAACTATTGGAGAAGTGC	1490
Db	1081	TTTGCAAAAATCATTTGAAGACCTAAGAAGGAAGGAAAGGAAAGTATTGGAGAAGTGC	1140
Qy	1491	CTTGTCCAGAGCCTGACAAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGGTC	1550
Db	1141	CTTGTCCAGAGCCTGACAAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGGTC	1200
Qy	1551	TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG	1610
Db	1201	TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG	1260
Qy	1611	CTGGCCTCCTTGAATCTCCTTCGGGTGTCAAGTGTCTTTGTGGCTATTGCAGTCAAAAGST	1670
Db	1261	CTGGCCTCCTTGAATCTCCTTCGGGTGTCAAGTGTCTTTGTGGCTATTGCAGTCAAAAGST	1320
Qy	1671	CTCAGCAATTCGAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTCCTCCAGGAGAGCCCT	1730
Db	1321	CTCAGCAATTCGAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTCCTCCAGGAGAGCCCT	1380
Qy	1731	GTTTTCATGTCCAGACATTTACAAAGACCCACGAAAGCTCTGGTCTTTGAGGAGGCCACC	1790
Db	1381	GTTTTCATGTCCAGACATTTACAAAGACCCACGAAAGCTCTGGTCTTTGAGGAGGCCACC	1440
Qy	1791	TTGTCTATGTCCAGACATTTACAAAGACCCACGAAAGCTCTGGTCTTTGAGGAGGGAAC	1850
Db	1441	TTGTCTATGTCCAGACATTTACAAAGACCCACGAAAGCTCTGGTCTTTGAGGAGGGAAC	1500
Qy	1851	GGGCATGCTTCTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGAGG	1910
Db	1501	GGGCATGCTTCTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGAGG	1560
Qy	1911	AACAGCTGGGCCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCTCAAGGGGATGATGTTA	1970
Db	1561	AACAGCTGGGCCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCTCAAGGGGATGATGTTA	1620
Qy	1971	GGGTCCTGGGCAACACGGGAGTGGTAAAGAGCAGCCTGTTGTCAGCCATCCTGGAGGAG	2030
Db	1621	GGGTCCTGGGCAACACGGGAGTGGTAAAGAGCAGCCTGTTGTCAGCCATCCTGGAGGAG	1680
Qy	2031	ATGCATCTGCTCGAGGGCTCGGTGGGGGTGCAGGGAAGCCCTGGCCTATGTCCCCCAGCAG	2090
Db	1681	ATGCATCTGCTCGAGGGCTCGGTGGGGGTGCAGGGAAGCCCTGGCCTATGTCCCCCAGCAG	1740
Qy	2091	GCCTGGATCGTCAGCGGGAACATCAGGGAGAACATCCTCATGGGAGCGCATATGACAAG	2150
Db	1741	GCCTGGATCGTCAGCGGGAACATCAGGGAGAACATCCTCATGGGAGCGCATATGACAAG	1800
Qy	2151	GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACCTTCTGCC	2210
Db	1801	GCCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGGACCTGGAACCTTCTGCC	1860
Qy	2211	TTTGGAGACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGGCAGAAACAG	2270
Db	1861	TTTGGAGACATGACAGAGATTGGAGAGCGGGGCCCTCAACCTCTCTGGGGGGCAGAAACAG	1920
Qy	2271	AGGATCAGCCTGGCCCGCCGCTCTATTCCGACCGTCAGATCTACCTGCTGGACGACCC	2330
Db	1921	AGGATCAGCCTGGCCCGCCGCTCTATTCCGACCGTCAGATCTACCTGCTGGACGACCC	1980
Qy	2331	CTGTCGTGTGGACGCGCCACGTTGGGAAGCACATTTTGGAGAGTGCATTAAAGAAGACA	2390
Db	1981	CTGTCGTGTGGACGCGCCACGTTGGGAAGCACATTTTGGAGAGTGCATTAAAGAAGACA	2040
Qy	2391	CTCAGGGGGAAGACGGTCTGCTCTGGTGACCCACCGCTGCAGTGCAGTACTTAGAATTTTGTG	2450
Db	2041	CTCAGGGGGAAGACGGTCTGCTCTGGTGACCCACCGCTGCAGTGCAGTACTTAGAATTTTGTG	2100

Qy	2451	CAGATCATTTTGTGGAAAAATGGGAAAAATCTGTGAAAAATGGAACCTCACAGTGAGTTAATG	2510
Db	2101	CAGATCATTTTGTGGAAAAATGGGAAAAATCTGTGAAAAATGGAACCTCACAGTGAGTTAATG	2160
Qy	2511	CAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCCGAC	2570
Db	2161	CAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTC----	2216
Qy	2571	ATGTTGCAGGACACAGCAAAAGATAGCAGAGAAGCCAAAGSTAGAAAAGTCAGGCTCTGGCC	2630
Db	2217	-----	2216
Qy	2631	ACCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCCGGAGCATCAGCTCACACAGGAG	2690
Db	2217	-----	2216
Qy	2691	GAGGAGTGAAGAAGGCTCCTTGAGTTGGAGGCTTACCACCACCTACATCCAGGCAGCT	2750
Db	2217	-----	2216
Qy	2751	GGAGTTACATGGTCTCTTTGCATAATTTTCTTCTTCGTGGTGCTGATCGTCTTCTTAACG	2810
Db	2217	-----	2216
Qy	2811	ATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGAGGGCTCGGGGACCAATAGCAGC	2870
Db	2217	-----	2216
Qy	2871	CGAGAGCAATGGAACCATGGCAGACCTGGGCAACAATTGCAGACAATCCTCAACTGTCC	2930
Db	2217	-----	2216
Qy	2931	TTCTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCC	2990
Db	2217	-----	2216
Qy	2991	TCAGGGATTTTCACCAAAGTCACGAGGAAGGCATCCACGSCCTGCACAACAAGCTCTTC	3050
Db	2217	-----	2216
Qy	3051	AACAAGGTTTCCGCTGCCCATGAGTTTCTTTTGACACCATCCCAATAGCCCGCTTTTG	3110
Db	2217	-----GGTTTCCGCTGCCCATGAGTTTCTTTTGACACCATCCCAATAGCCCGCTTTTG	2271
Qy	3111	AACTGCTTCGAGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCCATCTTTTCAGAGCAG	3170
Db	2272	AACTGCTTCGAGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCCATCTTTTCAGAGCAG	2331
Qy	3171	TTCTGGTCTGTCCTTAATGGTGATCGCCGTCCTGTTGATGTGCTGCTGCCA	3230
Db	2332	TTCTGGTCTGTCCTTAATGGTGATCGCCGTCCTGTTGATGTGCTGCTGCCA	2391
Qy	3231	TATATCCTGTTAATGGGAGCCATAATCATGGTTATTGCTTCATTTATATATGATGTTTC	3290
Db	2392	TATATCCTGTTAATGGGAGCCATAATCATGGTTATTGCTTCATTTATATATGATGTTTC	2451
Qy	3291	AAGAAAGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTCTCC	3350
Db	2452	AAGAAAGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTCTCC	2511
Qy	3351	CACATCCTCAATTCTCTGCAAGSCCTGAGCTCCATCCATGCTATATGAAAAAACTGAAGAC	3410
Db	2512	CACATCCTCAATTCTCTGCAAGSCCTGAGCTCCATCCATGCTATGAAAAAACTGAAGAC	2571
Qy	3411	TTTCATCAGCCAGTTTAAGAGGCTGACTGATGCGCAGAATAACTACCTGCTGTTGTTCTA	3470
Db	2572	TTTCATCAGCCAGTTTAAGAGGCTGACTGATGCGCAGAATAACTACCTGCTGTTGTTCTA	2631
Qy	3471	TCCTCCACACGATGGATGGCATGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT	3530
Db	2632	TCCTCCACACGATGGATGGCATGAGGCTGGAGATCATGACCAACCTTGTGACCTTGGCT	2691

QY 1071 TTTGCCCTTGAGAAGCTCATCCAATTTAAGTCTGTAAATACACATCACCTCAGGAGAGGCC 1130
Db 721 TTTGCCCTTGAGAAGCTCATCCAATTTAAGTCTGTAAATACACATCACCTCAGGAGAGGCC 780
QY 1131 ATCAGCTTCTTCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGCTGCTATGGACCCCTA 1190
Db 781 ATCAGCTTCTTCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGCTGCTATGGACCCCTA 840
QY 1191 GTACTGATCACCTGCGCATCGCTGGTCATCTGCAGCATTTCTTCTTACTTCAATTATTGGA 1250
Db 841 GTACTGATCACCTGCGCATCGCTGGTCATCTGCAGCATTTCTTCTTACTTCAATTATTGGA 900
QY 1251 TACACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTTTCCCACCTGGCGGTATTTCATG 1310
Db 901 TACACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTTTCCCACCTGGAGGTATTTCATG 960
QY 1311 ACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT 1370
Db 961 ACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT 1020
QY 1371 GTGACCAGTGAAGTTCTCACTTGCAATTAAAGCTGATTAATAATGTACACATGGGAGAAACCA 1430
Db 1021 GTGACCAGTGAAGTTCTCACTTGCAATTAAAGCTGATTAATAATGTACACATGGGAGAAACCA 1080
QY 1431 TTTGCAAAAATCATTTGAAGCCTTAAGAAGGAAGGAAAGAAACTATTGGAGAAGTGCGGG 1490
Db 1081 TTTGCAAAAATCATTTGAAGCCTTAAGAAGGAAGGAAAGAACTATTGGAGAAGTGCGGG 1140
QY 1491 CTTGTCCAGAGCCTGACAAAGTATAACCTTGTTTCATCATCCCCACAGTGGCCACAGCGGTC 1550
Db 1141 CTTGTCCAGAGCCTGACAAAGTATAACCTTGTTTCATCATCCCCACAGTGGCCACAGCGGTC 1200
QY 1551 TGGGTTCTCATCCACACATCCTTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG 1610
Db 1201 TGGGTTCTCATCCACACATCCTTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG 1260
QY 1611 CTGGCCTCCTTGAATCTCCTTCGGGCTGTCAGTGTCTTTGTGGCTATTGTCAGTCAAAGGT 1670
Db 1261 CTGGCCTCCTTGAATCTCCTTCGGGCTGTCAGTGTCTTTGTGGCTATTGTCAGTCAAAGGT 1320
QY 1671 CTCAGGAATTCCAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTTCTCCAGGAGAGCCCT 1730
Db 1321 CTCAGGAATTCCAAGTCTGCAGTGATGAGGTTCAAGAAAGTTTTTCTCCAGGAGAGCCCT 1380
QY 1731 GTTTTCTATGTCCAGACATTTACAAGACCCAGCAAAAGCTCTGGTCTTTTGAGGAGGCCACC 1790
Db 1381 GTTTTCTATGTCCAGACATTTACAAGACCCAGCAAAAGCTCTGGTCTTTTGAGGAGGCCACC 1440
QY 1791 TTGTCTATGGCAACAGACCTGTCCCGGGATCGTCAATGGGGCACTGGAGCTGGAGAGGAAC 1850
Db 1441 TTGTCTATGGCAACAGACCTGTCCCGGGATCGTCAATGGGGCACTGGAGCTGGAGAGGAAC 1500
QY 1851 GGGCATGCTTCTGAGGGGATGACCAGGCTTAGAGATGCCCTCGGGCCAGAGGAAGGG 1910
Db 1501 GGGCATGCTTCTGAGGGGATGACCAGGCTTAGAGATGCCCTCGGGCCAGAGGAAGGG 1560
QY 1911 AACAGCCTGSGCCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATGTTA 1970
Db 1561 AACAGCCTGSGCCCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATGTTA 1620
QY 1971 GGGGTCTGCGGAACACGGGGAGTGGTAAGAGCAGCCCTGTTGTACGCCATCCTGGAGGAG 2030
Db 1621 GGGGTCTGCGGAACACGGGGAGTGGTAAGAGCAGCCCTGTTGTACGCCATCCTGGAGGAG 1680
QY 2031 ATGCATTTGCTCAGGGGCTCGGTGGGGTGCAGGGAAGCCTGGCCTATGTCCTCCAGCAG 2090
Db 1681 ATGCATTTGCTCAGGGGCTCGGTGGGGTGCAGGGAAGCCTGGCCTATGTCCTCCAGCAG 1740
QY 2091 GCCTGGATCGTCAGCGGGAACATCAGGAGAAACATCCTCATGGAGGCGCATATGACAAG 2150
Db 1741 GCCTGGATCGTCAGCGGGAACATCAGGAGAAACATCCTCATGGAGGCGCATATGACAAG 1800
QY 2151 GCCCGATACCCAGGTTGCTCCACTGCTGCCCTGAAATCGGGACCTTGGAACTTCTGCC 2210

Db 1801 GCCGATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGACCTTGAACCTTCTGCC 1860
QY 2211 TTTGGAGACATGACAGAGATTGGAGAGCGGGCCTCAACCTCTCTGGGGGCGAGAAACAG 2270
Db 1861 TTTGGAGACATGACAGAGATTGGAGAGCGGGCCTCAACCTCTCTGGGGGCGAGAAACAG 1920
QY 2271 AGGATCAGCCTGGCCCCGGCCGTCTATTTCCGACCGTCAGATCTACCTGCTGGACGACCCC 2330
Db 1921 AGGATCAGCCTGGCCCCGGCCGTCTATTTCCGACCGTCAGATCTACCTGCTGGACGACCCC 1980
QY 2331 CTGCTGCTGTGGACGCCACCGTGGGGAAGCACATTTTGGAGAGTGCATTTAAGAAAGACA 2390
Db 1981 CTGCTGCTGTGGACGCCACCGTGGGGAAGCACATTTTGGAGAGTGCATTTAAGAAAGACA 2040
QY 2391 CTCAGGGGAAGACGGTCTGCTGCTGACCCACCGTGCAGTGCAGTACTTAGAATTTTGTGGC 2450
Db 2041 CTCAGGGGAAGACGGTCTGCTGCTGACCCACCGTGCAGTGCAGTACTTAGAATTTTGTGGC 2100
QY 2451 CAGATCATTTTGTGGAAAAATGGAAAAATCTGTGAAAAATGGAACCTCACAGTGAGTTAATG 2510
Db 2101 CAGATCATTTTGTGGAAAAATGGAAAAATCTGTGAAAAATGGAACCTCACAGTGAGTTAATG 2160
QY 2511 CAGAAAAAGGGGAAATATGCCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGGAC 2570
Db 2161 CAGAAAAAGGGGAAATATGCCCCAACTTATCCAGAAGATGCACAAGGAAGCCACTTCGGAC 2220
QY 2571 ATGTTGCAGGACACAGCAAAAGATAGCAGAGAAGCCAAAGGTAGAAAAGTCAGGCTCTGGCC 2630
Db 2221 ATGTTGCAGGACACAGCAAAAGATAGCAGAGAAGCCAAAGGTAGAAAAGTCAGGCTCTGGCC 2280
QY 2631 ACCTCCTGGAAGAGTCTCTCAAGGGAAATGCTGTGCCGGAGCATCAGCTCACACAGGAG 2690
Db 2281 ACCTCCTGGAAGAGTCTCTCAAGGGAAATGCTGTGCCGGAGCATCAGCTCACACAGGAG 2340
QY 2691 GAGGAGATGGAAGAAGGCTCCTTGAGTTGGAGGCTTACCAACACTACATCCAGGCAGCT 2750
Db 2341 GAGGAGATGGAAGAAGGCTCCTTGAGTTGGAGGCTTACCAACACTACATCCAGGCAGCT 2400
QY 2751 GGAGGTTACATGGTCTCTTGCAATAATTTCTTCTCGTGGTGTGATCGTCTTCTTAACG 2810
Db 2401 GGAGGTTACATGGTCTCTTGCAATAATTTCTTCTCGTGGTGTGATCGTCTTCTTAACG 2460
QY 2811 ATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGCAGC 2870
Db 2461 ATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGCAGC 2520
QY 2871 CGAGAGAGCAATGGAACCATGSCAGACCTGGGCAACATTTGCAGACAATCCTCAACTGTCC 2930
Db 2521 CGAGAGAGCAATGGAACCATGSCAGACCTGGGCAACATTTGCAGACAATCCTCAACTGTCC 2580
QY 2931 TTCTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGGTCTGCTCC 2990
Db 2581 TTCTACCAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGGTCTGCTCC 2640
QY 2991 TCAGGATTTTCAACCAAGTACGAGGAAGGCATCCAGGCCCTGCACAACAAGCTCTTTC 3050
Db 2641 TCAGGATTTTCAACCAAGTACGAGGAAGGCATCCAGGCCCTGCACAACAAGCTCTTTC 2700
QY 3051 AACAAAGTTTTCGGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCCCGGCTTTTG 3110
Db 2701 AACAAAGTTTTCGGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCCCGGCTTTTG 2760
QY 3111 AACTGCTTCGCAGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAGCAG 3170
Db 2761 AACTGCTTCGCAGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCATCTTTTCAGAGCAG 2820
QY 3171 TTCTTGGTCTCTTCCTTAATGGTGATCGCCCGTCCCTGTTGATGTGTCAGTGTCTCTCCA 3230
Db 2821 TTCTTGGTCTCTTCCTTAATGGTGATCGCCCGTCCCTGTTGATGTGTCAGTGTCTCTCCA 2880
QY 3231 TATATCCTGTTAATGGGAGCCCAATCATGGTTATTTGCTTCATTTATATATATGATGTTTC 3290

Db	2881	TATATCCTGTTAATGGGAGCCATAATCATGGTTATTGCTTCAATTTATTATATGATGTTT	2940
QY	3291	AAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTCTCC	3350
Db	2941	AAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTCTCC	3000
QY	3351	CACATCCTCAATTCTGCAAGGCCTGAGCTCCATCCCATGTCTATGGA AAAA ACTGAAGAC	3410
Db	3001	CACATCCTCAATTCTGCAAGGCCTGAGCTCCATCCCATGTCTATGGA AAAA ACTGAAGAC	3060
QY	3411	TTCATCAGCCAGT 3423	
Db	3061	TTCATCAGCCAGT 3073	
RESULT 13			
AX135180			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 2930; Conservative			
QY	351	ATGACTAGGAAGAGACATACTGGGTGCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	410
Db	1	ATGACTAGGAAGAGACATACTGGGTGCCAACTCTTCTGGTGGCCTCGTGAATCGTGGC	60
QY	411	ATCGACATAGGCGATGACATGTTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT	470
Db	61	ATCGACATAGGCGATGACATGTTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT	120
QY	471	GGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGCAGCTGTCCCACCG	530
Db	121	GGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAGGGAGGGCAGCTGTCCCACCG	180
QY	531	TGGGGGAAGTATGATGCTGCCTTGAGAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT	590
Db	181	TGGGGGAAGTATGATGCTGCCTTGAGAACCATGATTCCTTCCGTCCCAAGCCGAGGTTT	240
QY	591	CCTGCCCCCAGCCCTGGACAAATGCTGGCTGTTCTCCTACCTCACCGTGTATGGCTC	650
Db	241	CCTGCCCCCAGCCCTGGACAAATGCTGGCTGTTCTCCTACCTCACCGTGTATGGCTC	300
QY	651	ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTG	710
Db	301	ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAACACCATCCCTCCACTG	360
QY	711	TCAGTCCATGATGCCTCAGACAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA	770
Db	361	TCAGTCCATGATGCCTCAGACAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAAGAA	420
QY	771	GTCTCAAGGGCAGGGATTGAAAAAGCTTCAGTGTCTTCTGGTATGCTGAGGTTCCAGAGA	830

Db	421	GTCTCAAGCGGAGGATTGAAAAAGCTTCAGTGTCTGGTGATGCTGAGGTTCCAGAGA	480
QY	831	ACAAGGTTGATTTTCGATGCACATTCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGGG	890
Db	481	ACAAGGTTGATTTTCGATGCACATTCTGGGCATCTGCTTCTGCATTGCCAGTGTACTCGGG	540
QY	891	CCAATATTGATTATACCAAAAGATCCTGGAAATATTCAGAAAGACAGTTGGGCAATGTTGTC	950
Db	541	CCAATATTGATTATACCAAAAGATCCTGGAAATATTCAGAAAGACAGTTGGGCAATGTTGTC	600
QY	951	CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTCAGTTTC	1010
Db	601	CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTCAGTTTC	560
QY	1011	TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTTCCGAGCAGCTGTTTCTCTCC	1070
Db	661	TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTTCCAAGCAGCTGTTTCTCTCC	720
QY	1071	TTTGCCCTTTGAGAAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC	1130
Db	721	TTTGCCCTTTGAGAAAGCTCATCCAAATTTAAGTCTGTAATACACATCACCTCAGGAGAGGCC	780
QY	1131	ATCAGCTTCTTCACCCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTA	1190
Db	781	ATCAGCTTCTTCACCCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTA	840
QY	1191	GTA CTGATCACCTGCGCATCGCTGGTCACTCGAGCATTTCTTCCCTACTTCATTATTGGA	1250
Db	841	GTA CTGATCACCTGCGCATCGCTGGTCACTCGAGCATTTCTTCCCTACTTCATTATTGGA	900
QY	1251	TACACTGCATTTAFTGCCATCTTATGCTATCTCCTGGTTTTCCCACCTGGCGGTATTCTATG	1310
Db	901	TACACTGCATTTAFTGCCATCTTATGCTATCTCCTGGTTTTCCCACCTGGAGGTATTCTATG	960
QY	1311	ACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCCAGCCCATCCGT	1370
Db	961	ACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCCAGCCCATCCGT	1020
QY	1371	GTGACCAGTGAAGTTCTCACATTGCAATTAAAGTGAATTAATAATGTACACATGGGAGAAACCA	1430
Db	1021	GTGACCAGTGAAGTTCTCACATTGCAATTAAAGTGAATTAATAATGTACACATGGGAGAAACCA	1080
QY	1431	TTTGC AAAAATCATTTGAAGACCTTAAGAAAGGAGGAAAGGAAACATA TTGAGAAAGTGCGGG	1490
Db	1081	TTTGC AAAAATCATTTGAAGACCTTAAGAAAGGAGGAAAGGAAAGCTATTGGAGAAAGTGCGGG	1140
QY	1491	CTTGTCCAGAGCCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGGTC	1550
Db	1141	CTTGTCCAGAGCCTGACAAGTATAACCTTGTTCATCATCCCCACAGTGGCCACAGCGGTC	1200
QY	1551	TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG	1610
Db	1201	TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCACAGCGTCAATGGCCTTCAGCATG	1260
QY	1611	CTGGCCTCCTTGAATCTCCTTCGGCTGTCAAGTGTCTTTTGTGCCCTATTGCAGTCAAAGGT	1670
Db	1261	CTGGCCTCCTTGAATCTCCTTCGGCTGTCAAGTGTCTTTTGTGCCCTATTGCAGTCAAAGGT	1320
QY	1671	CTCAGCAATTCCAAGTCTGCAGTGATGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCT	1730
Db	1321	CTCAGCAATTCCAAGTCTGCAGTGATGAGGTTCAAGAAGTTTTCCTCCAGGAGAGCCCT	1380
QY	1731	GTTTTCATGTCCAGACATTACAAGACCCACAGCAAGGCTCTGGTCTTTTGAGGAGGCCACC	1790
Db	1381	GTTTTCATGTCCAGACATTACAAGACCCACAGCAAGGCTCTGGTCTTTTGAGGAGGCCACC	1440
QY	1791	TTGT CATGGCAACAGACCTGTCCCGGGATCGTCAATGGGCACCTGGAGCTGGAGAGAAC	1850
Db	1441	TTGT CATGGCAACAGACCTGTCCCGGGATCGTCAATGGGCACCTGGAGCTGGAGAGAAC	1500
QY	1851	GGGCATGCTTCTGAGGGGATGACCAGGCCCTAGAGATGCCCTCGGGCCAGAGGAAGGG	1910

Db 1501 GGGCATGCTTCTGAGGGGATGACCAAGCCCTAGAGATGCCCTCGGGCCAGAGGAAGAGG 1560

QY 1911 AACAGCCTGGGCCCCAGAGTTGCACAAGATCAACCTGGTGTCTCCAAGGGGATGATGTTA 1970

Db 1561 AACAGCCTGGGCCCCAGAGTTGCACAAGATCAACCTGGTGTCTCCAAGGGGATGATGTTA 1620

QY 1971 GGGGTCTGGGCAACACGGGGAGTGGTAAGAGCAGCCCTGTTGTCAGCCCATCTTGAGGAG 2030

Db 1621 GGGGTCTGGGCAACACGGGGAGTGGTAAGAGCAGCCCTGTTGTCAGCCCATCTTGAGGAG 1680

QY 2031 ATGCACCTGCTCAGGGGCTCGGTGGGGTGCAGGGAAGCCCTGGCCTATGTCCCCCAGCAG 2090

Db 1681 ATGCACCTGCTCAGGGGCTCGGTGGGGTGCAGGGAAGCCCTGGCCTATGTCCCCCAGCAG 1740

QY 2091 GCCTGGATCGTCAGCGGGAACATCAGGGGAGAACATCCTCATGSGAGGCGCATATGACAAG 2150

Db 1741 GCCTGGATCGTCAGCGGGAACATCAGGGGAGAACATCCTCATGSGAGGCGCATATGACAAG 1800

QY 2151 GCCCCATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGAGCCCTGGAACTTCTGCC 2210

Db 1801 GCCCCATACCTCCAGGTGCTCCACTGCTGCTCCCTGAATCGGAGCCCTGGAACTTCTGCC 1860

QY 2211 TTTGGAGACATGACAGAGATTGGAGACGGGGCCCTCAACCTCTCTGGGGGGCAGAAACAG 2270

Db 1861 TTTGGAGACATGACAGAGATTGGAGACGGGGCCCTCAACCTCTCTGGGGGGCAGAAACAG 1920

QY 2271 AGGATCAGCTGGCCCCGCGCGCTCTATTCCGACCGTCAGATCCTGTGGACGACCCC 2330

Db 1921 AGGATCAGCTGGCCCCGCGCGCTCTATTCCGACCGTCAGATCCTGTGGACGACCCC 1980

QY 2331 CTGTCTGCTGTGGACGCCCCACGTGGGGAAGCACATTTTGTAGSAGTGCATTAAAGAAGACA 2390

Db 1981 CTGTCTGCTGTGGACGCCCCACGTGGGGAAGCACATTTTGTAGSAGTGCATTAAAGAAGACA 2040

QY 2391 CTCAGSGGGAAGACGGTCGTCCTGCTGACCCACCAGCTGCAGTACTTAGAATTTTGTGGC 2450

Db 2041 CTCAGSGGGAAGACGGTCGTCCTGCTGACCCACCAGCTGCAGTACTTAGAATTTTGTGGC 2100

QY 2451 CAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACTCACAGTGAGTTAATG 2510

Db 2101 CAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACTCACAGTGAGTTAATG 2160

QY 2511 CAGAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGAGCCACTTCGGAC 2570

Db 2161 CAGAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGAGCCACTTCGGAC 2220

QY 2571 ATGTTGAGGACACAGCAAAGATAGCAGAGAAGCCAAAGGTAGAAAGTCAGGCTCTGGCC 2630

Db 2221 ATGTTGAGGACACAGCAAAGATAGCAGAGAAGCCAAAGGTAGAAAGTCAGGCTCTGGCC 2280

QY 2631 ACCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCCGGAGCATCAGCTCACACAGGAG 2690

Db 2281 ACCTCCCTGGAAGAGTCTCTCAACGGAAATGCTGTGCCGGAGCATCAGCTCACACAGGAG 2340

QY 2691 GAGGAGATGGAAGAAGGCTCCCTTGAGTTGGAGGGTCTACCAACCACTACATCCAGGCAGCT 2750

Db 2341 GAGGAGATGGAAGAAGGCTCCCTTGAGTTGGAGGGTCTACCAACCACTACATCCAGGCAGCT 2400

QY 2751 GGAGGTTACATGGTCTCTTGCATAATTTTCTTCTCGTGGTGCTGATCGTCTTCTTAACG 2810

Db 2401 GGAGGTTACATGGTCTCTTGCATAATTTTCTTCTCGTGGTGCTGATCGTCTTCTTAACG 2460

QY 2811 ATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGCAGC 2870

Db 2461 ATCTTCAGCTTCTGGTGGCTGAGCTACTGGTTGGAGCAGGGCTCGGGGACCAATAGCAGC 2520

QY 2871 CGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTGCAGACAATCCTCAACTGTCC 2930

Db 2521 CGAGAGCAATGGAACCATGGCAGACCTGGGCAACATTGCAGACAATCCTCAACTGTCC 2580

QY 2931 TTCTACAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCC 2990

Db 2581 TTCTACAGCTGGTGTACGGGCTCAACGCCCTGCTCCTCATCTGTGTGGGGTCTGCTCC 2640

QY 2991 TCAGGGATTTTCACCAAAGTCACGAGGAAGGCATCCACGGSCCTGCACAACAAGCTCTTC 3050

Db 2641 TCAGGGATTTTCACCAAAGTCACGAGGAAGGCATCCACGGSCCTGCACAACAAGCTCTTC 2700

QY 3051 AACAAAGTTTCCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCCGGCTTTTG 3110

Db 2701 AACAAAGTTTCCGCTGCCCATGAGTTTCTTTGACACCATCCCAATAGCCGGCTTTTG 2760

QY 3111 AACTGCTTCGACGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCCATCTTTTCAGAGCAG 3170

Db 2761 AACTGCTTCGACGGGGACTTGGAAACAGCTGGACCAGCTCTTGCCCCATCTTTTCAGAGCAG 2820

QY 3171 TTCCTGGTCTCTGCTCTTAATGGTGATCGCCGCTCTGTTGATTGTCAAGTGTCTCCA 3230

Db 2821 TTCCTGGTCTCTGCTCTTAATGGTGATCGCCGCTCTGTTGATTGTCAAGTGTCTCCA 2880

QY 3231 TATATCCTGTTAATGGGAGCCATAATCATGGTTATTGCTTCATTTATTATATG 3284

Db 2881 TATATCCTGTTAATGGGAGCCATAATCATGGTTATTGCTTCATTTATTATATG 2934

RESULT 14
AX135186

LOCUS AX135186 2448 bp DNA linear PAT 29-MAY-2001
DEFINITION Sequence 19 from Patent WO0132706.
ACCESSION AX135186

VERSION AX135186.1 GI:14271538

KEYWORDS

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2448)

REFERENCE

AUTHORS Haras,M., Donoho,G., Turner,C.A., Nehls,M., Friedrich,G.,
Zambrowicz,B. and Sands,A.T.

TITLE Novel human transporter proteins and polynucleotides encoding the
same

JOURNAL

Patent: WO 0132706-A 19 10-MAY-2001;
Lexicon Genetics Incorporated (US)

FEATURES

Location/Qualifiers

Source

1..2448

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 582 a 619 c 639 g 608 t

ORIGIN

Query Match 45.5%; Score 2213.2; DB 6; Length 2448;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2215; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 351 ATGACTAGGAAGAGGACATACTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGC 410

Db 1 ATGACTAGGAAGAGGACATACTGGTGCCCAACTCTTCTGGTGGCCTCGTGAATCGTGC 60

QY 411 ATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT 470

Db 61 ATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAAAACCTATACTCTCCAAGAT 120

QY 471 GGCCCTGGAGTCAGCAAGAGAGAAAATCCTGAGGCTCCAGGAGGGCAGCTGTCCCACCG 530

Db 121 GGCCCTGGAGTCAGCAAGAGAGAAAATCCTGAGGCTCCAGGAGGGCAGCTGTCCCACCG 180

QY 531 TGGGGGAAGTATGATGCTGCCCTTGAGAACCATGATTCCTTCCTCCCAAGCCGAGGTTT 590

Db 181 TGGGGGAAGTATGATGCTGCCCTTGAGAACCATGATTCCTTCCTCCCAAGCCGAGGTTT 240

QY 591 CCTGCCCCCAGCCCTGGACAATGCTGGCCTGTTCTCTACCTCACCGTGTCAATGGCTC 650

Db 241 CCTGCCCCCAGCCCTGGACAATGCTGGCCTGTTCTCTACCTCACCGTGTCAATGGCTC 300

QY 651 ACCCGGCTCATGATCCAAAGCTTACGGAGTCGTTAGATGAGAACACCATCCCTCCACTG 710

|||||

Db	301	ACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGATGAGAAACCATCCCTCCACTG	360
Qy	711	TCAGTCCATGATGCCCTCAGACAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAGAA	770
Db	361	TCAGTCCATGATGCCCTCAGACAAAAATGTCCAAAGGCTTCACCGCCTTTGGGAAGAGAA	420
Qy	771	GTCTCAAGGCGAGGGATTGAAAAGCTTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGA	830
Db	421	GTCTCAAGGCGAGGGATTGAAAAGCTTTCAGTGCTTCTGGTGATGCTGAGGTTCCAGAGA	480
Qy	831	ACAAGGTTGATTTTCGATGCACTTCTGGGCATCTGCTTCGCATTCGCAGTGTACTCGGG	890
Db	481	ACAAGGTTGATTTTCGATGCACTTCTGGGCATCTGCTTCGCATTCGCAGTGTACTCGGG	540
Qy	891	CCAATATTGATTATACCAAAAGATCCTTGGAAATATTCAGAAAGACAGTTGGGGAATGTGTC	950
Db	541	CCAATATTGATTATACCAAAAGATCCTTGGAAATATTCAGAAAGACAGTTGGGGAATGTGTC	600
Qy	951	CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC	1010
Db	601	CATGGAGTGGGACTCTGCTTTGCCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTC	660
Qy	1011	TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCC	1070
Db	661	TCCTCCAGTTGGATCATCAACCAACGCACAGCCATCAGGTTCCAAGCAGCTGTTTCCTCC	720
Qy	1071	TTTGGCCTTTGAGAAAGCTCATCCAATTTAAGTCTGTAAATACACATCACCTCAGGAGAGGCC	1130
Db	721	TTTGGCCTTTGAGAAAGCTCATCCAATTTAAGTCTGTAAATACACATCACCTCAGGAGAGGCC	780
Qy	1131	ATCAGCTTCTTACCCGGTGATGTAAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTA	1190
Db	781	ATCAGCTTCTTACCCGGTGATGTAAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTA	840
Qy	1191	GTA CTGATCACCTGCGCATCGTGGTTCATCTGCAGCATTTCTTCTCTACTTTCATTATGGA	1250
Db	841	GTA CTGATCACCTGCGCATCGTGGTTCATCTGCAGCATTTCTTCTCTACTTTCATTATGGA	900
Qy	1251	TACACTGCATTTATTGGCCATCTTATGCTATCTCCTGGTTTTCCTCCACTGGCGGTATTCATG	1310
Db	901	TACACTGCATTTATTGGCCATCTTATGCTATCTCCTGGTTTTCCTCCACTGGAGGTATTCATG	960
Qy	1311	ACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT	1370
Db	961	ACAAGAAATGGCTGTGAAGGCTCAGCATCACACATCTGAGGTCAGCGACCGCATCCGT	1020
Qy	1371	GTGACCAAGTGAAGTTCTCACTTGCATTAAAGCTGATTAAATGTACACATGGGAGAAACCA	1430
Db	1021	GTGACCAAGTGAAGTTCTCACTTGCATTAAAGCTGATTAAATGTACACATGGGAGAAACCA	1080
Qy	1431	TTTGCAAAAAATCATTTGAAGACCTAAGAAGGAAGGAAAGGAAAACTATTGGAGAAGTCGGGG	1490
Db	1081	TTTGCAAAAAATCATTTGAAGACCTAAGAAGGAAGGAAAGGAAAGCTATTGGAGAAGTCGGGG	1140
Qy	1491	CTTGCTCCAGAGCCTTGACAAGTATAACCTTTGTTTCATCATCCCCACAGTGGCCACAGCGGTC	1550
Db	1141	CTTGCTCCAGAGCCTTGACAAGTATAACCTTTGTTTCATCATCCCCACAGTGGCCACAGCGGTC	1200
Qy	1551	TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCAGACGGTCAATGGCCTTTCAGCATG	1610
Db	1201	TGGGTTCTCATCCACACATCCTTAAAGCTGAAACTCAGACGGTCAATGGCCTTTCAGCATG	1260
Qy	1611	CTGGCCTCCTTGAATCTCCTTCGGCTGTCAAGTGTTCAGTGTTCCTTGTGCCTATTGCAGTCAAAGGT	1670
Db	1261	CTGGCCTCCTTGAATCTCCTTCGGCTGTCAAGTGTTCAGTGTTCCTTGTGCCTATTGCAGTCAAAGGT	1320
Qy	1671	CTCAGCAATTCCAAGTCTCGAGTGATGAGGTTCAAGAAGTTTTTTCTCCAGGAGAGCCCT	1730
Db	1321	CTCAGCAATTCCAAGTCTCGAGTGATGAGGTTCAAGAAGTTTTTTCTCCAGGAGAGCCCT	1380
Qy	1731	GTTTTCTATGTCCAGACATTACAAGACCCCGCAGCAAGCTCTGGTCTTTTGAGGAGGCCACC	1790
Db	1381	GTTTTCTATGTCCAGACATTACAAGACCCCGCAGCAAGCTCTGGTCTTTTGAGGAGGCCACC	1440

Qy	1791	TTGTTCATGGCAACAGACACCTGTCCCGGGATCGTCAATGGGCGACTGGAGCTGGAGAGGAAC	1850
Db	1441	TTGTTCATGGCAACAGACACCTGTCCCGGGATCGTCAATGGGCGACTGGAGCTGGAGAGGAAC	1500
Qy	1851	GGGCATGCTTCTGAGGGGATGACCAGGCCTTAGAGATGCCTTCGGGCCAGAGGAAGAAGG	1910
Db	1501	GGGCATGCTTCTGAGGGGATGACCAGGCCTTAGAGATGCCTTCGGGCCAGAGGAAGAAGG	1560
Qy	1911	AACAGCCTGGGCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATGTTA	1970
Db	1561	AACAGCCTGGGCCAGAGTTGCACAAGATCAACCTGGTGGTGTCCAAGGGGATGATGTTA	1620
Qy	1971	GGGTCTCGGCAACACAGGGGAGTGGTAAGAGCAGCCCTGTGTTCAGCCATCCTGGAGGAG	2030
Db	1621	GGGTCTCGGCAACACAGGGGAGTGGTAAGAGCAGCCCTGTGTTCAGCCATCCTGGAGGAG	1680
Qy	2031	ATGCACCTTCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCTGGCCTATGTCCCCCAGCAG	2090
Db	1681	ATGCACCTTCTCGAGGGCTCGGTGGGGTGCAGGGAAGCCTGGCCTATGTCCCCCAGCAG	1740
Qy	2091	GCCTGGATCGTCAGCGGGAACATCAGGGGAGAACATCCTCATCGGAGGCGCATATGACAAG	2150
Db	1741	GCCTGGATCGTCAGCGGGAACATCAGGGGAGAACATCCTCATCGGAGGCGCATATGACAAG	1800
Qy	2151	GCCCGATACCTCCAGGTGCTCCACTGCTGCCCTGAATCGGGACCTTGGAACTTCTGCCC	2210
Db	1801	GCCCGATACCTCCAGGTGCTCCACTGCTGCCCTGAATCGGGACCTTGGAACTTCTGCCC	1860
Qy	2211	TTTGAGAGACATGACAGAGATTGGAGAGCGGGCCTCAACCTCTCTGGGGGCAGAAACAG	2270
Db	1861	TTTGAGAGACATGACAGAGATTGGAGAGCGGGCCTCAACCTCTCTGGGGGCAGAAACAG	1920
Qy	2271	AGGATCAGCCTGGCCCGCGCGTCTATTCCGACCCGTACAGTCTACCTGCTGGACGACCCC	2330
Db	1921	AGGATCAGCCTGGCCCGCGCGTCTATTCCGACCCGTACAGTCTACCTGCTGGACGACCCC	1980
Qy	2331	CTGCTCTGCTGTGGACGCCACCTGGGGAAGCACATTTTTCAGGAGTGCATTAAAGAAGACA	2390
Db	1981	CTGCTCTGCTGTGGACGCCACCTGGGGAAGCACATTTTTCAGGAGTGCATTAAAGAAGACA	2040
Qy	2391	CTCAGGGGGGAAGACGGTCTGCTGGTGACCCACCAGCTGCAGTACTTAGAAATTTTGTGGC	2450
Db	2041	CTCAGGGGGGAAGACGGTCTGCTGGTGACCCACCAGCTGCAGTACTTAGAAATTTTGTGGC	2100
Qy	2451	CAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTAATG	2510
Db	2101	CAGATCATTTTGTGGAAAATGGGAAAATCTGTGAAAATGGAACCTCACAGTGAGTTAATG	2160
Qy	2511	CAGAAAAAGGGGAAATATGCCCAAACCTTATCCAGAAGATGCACAAGGAGCCACTTCG	2568
Db	2161	CAGAAAAAGGGGAAATATGCCCAAACCTTATCCAGAAGATGCACAAGGAGCCACTTCG	2218
RESULT 15			
AXI35188			
LOCUS			
DEFINITION			
AXI35188			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			

AXI35188 Sequence 21 from Patent WO0132706. linear PAT 29-MAY-2001

AXI35188

AXI35188.1 G1:14271539

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2586)

Harras,M., Donoho,G., Turner,C.A., Nehls,M., Friedrich,G., Zambrowicz,B. and Sands,A.T.

Novel human transporter proteins and polynucleotides encoding the same

Patent: WO 0132706-A 21 10-MAY-2001;

Lexicon Genetics incorporated (US)

Location/Qualifiers

source	1. .2586										
	/organism="Homo sapiens"										
	/db_xref="taxon:9606"										
BASE COUNT	619 a	656 c	667 g	644 t							
ORIGIN											
Query Match	45.5%;		Score 2213.2;	DB 6;	Length 2586;						
Best Local Similarity	99.9%;		Pred. No. 0;								
Matches 2215;	Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0;			
QY	351	ATGACTAGGAAGAGGACAT	ACTGGGTGCCCCAACTCT	TTCTGGTGGCCTCGTGAAT	CGTGGC	410					
Db	1	ATGACTAGGAAGAGGACAT	ACTGGGTGCCCCAACTCT	TTCTGGTGGCCTCGTGAAT	CGTGGC	60					
QY	411	ATCGACATAGGCGATGACAT	GGTTTTCAGGACTTATTTA	TATAAAACCTATACTCT	CCAAAGAT	470					
Db	61	ATCGACATAGGCGATGACAT	GGTTTTCAGGACTTATTTA	TATAAAACCTATACTCT	CCAAAGAT	120					
QY	471	GGCCCCCTGGAGTCAGCA	AGAGAGAAATCCTGAGGCT	CCAGGGAGGGCAGCTGT	CCCCACCG	530					
Db	121	GGCCCCCTGGAGTCAGCA	AGAGAGAAATCCTGAGGCT	CCAGGGAGGGCAGCTGT	CCCCACCG	180					
QY	531	TGGGGAAGTATGATGCT	GCCTTGAGAACCATGAT	TCCCTCCGTCCCAAGCCGAGGTTT	590						
Db	181	TGGGGAAGTATGATGCT	GCCTTGAGAACCATGAT	TCCCTCCGTCCCAAGCCGAGGTTT	240						
QY	591	CCTGCCCCCGACGCCCT	TGGACAATGCTGGCCTGT	TCTCCTACCTCACCGTGT	CATGGCTC	650					
Db	241	CCTGCCCCCGACGCCCT	TGGACAATGCTGGCCTGT	TCTCCTACCTCACCGTGT	CATGGCTC	300					
QY	651	ACCCGCTCATGATCCAA	AGCTTACGGAGTCGCTTAGAT	GAGAACACCATCCCTCCACTG	710						
Db	301	ACCCGCTCATGATCCAA	AGCTTACGGAGTCGCTTAGAT	GAGAACACCATCCCTCCACTG	360						
QY	711	TCAGTCCATGATGCCTC	AGACAAAATGTCCAAAGGCT	TACCGCCTTTGGGAAGAAGAA	770						
Db	361	TCAGTCCATGATGCCTC	AGACAAAATGTCCAAAGGCT	TACCGCCTTTGGGAAGAAGAA	420						
QY	771	GTCTCAAGGGCAGGGAT	TGAAAAGCTTCAGTGC	TTCTGGTGATGCTGAGGTTCCAGAGA	830						
Db	421	GTCTCAAGGGCAGGGAT	TGAAAAGCTTCAGTGC	TTCTGGTGATGCTGAGGTTCCAGAGA	480						
QY	831	ACAAGGTTGATTTTCGAT	GCACCTTCTGGGCATCTGCT	TCTGCATTGCCAGTGTACTCGGG	890						
Db	481	ACAAGGTTGATTTTCGAT	GCACCTTCTGGGCATCTGCT	TCTGCATTGCCAGTGTACTCGGG	540						
QY	891	CCAATATTGATATACCA	AGATCCTGGAATATTCA	GAAAGAGCAGTTGGGGAATGTTGTC	950						
Db	541	CCAATATTGATATACCA	AGATCCTGGAATATTCA	GAAAGAGCAGTTGGGGAATGTTGTC	600						
QY	951	CATGGAGTGGGACTCTG	CTTTTGCCCTTTTCTCT	CCGAATGTGTGAAGTCTCTGAGTTTC	1010						
Db	601	CATGGAGTGGGACTCTG	CTTTTGCCCTTTTCTCT	CCGAATGTGTGAAGTCTCTGAGTTTC	660						
QY	1011	TCCTCCAGTTGGATCAT	CAACCAACGCACAGCCAT	CAGGTTCCGAGCAGCTGTTTCCCTCC	1070						
Db	661	TCCTCCAGTTGGATCAT	CAACCAACGCACAGCCAT	CAGGTTCCGAGCAGCTGTTTCCCTCC	720						
QY	1071	TTTGGCCTTTGAGAAGCT	CATCCAATTTAAGTCTGT	TAATACACATCACCTCAGGAGAGGCC	1130						
Db	721	TTTGGCCTTTGAGAAGCT	CATCCAATTTAAGTCTGT	TAATACACATCACCTCAGGAGAGGCC	780						
QY	1131	ATCAGCTTCTTCACCGGT	GATGTAAACTACCTGT	TTGAAGGGGTGCTATGGACCCCTA	1190						
Db	781	ATCAGCTTCTTCACCGGT	GATGTAAACTACCTGT	TTGAAGGGGTGCTATGGACCCCTA	840						
QY	1191	GTACTGATCACCTGCGCAT	CGCTGGTCACTGT	GCAGCATTTCTTCCCTACTTCAATATGGA	1250						
Db	841	GTACTGATCACCTGCGCAT	CGCTGGTCACTGT	GCAGCATTTCTTCCCTACTTCAATATGGA	900						
QY	1251	TACACTGCATTTATTGCC	ATCTTATGCTATCTCT	CTGGTTTTCCTCCACTGGCGGTATTCATG	1310						
Db	901	TACACTGCATTTATTGCC	ATCTTATGCTATCTCT	CTGGTTTTCCTCCACTGGAGGTATTCATG	960						

QY	1311	ACAAGAATGGCTGTGA	AGGCTCAGCATCACAC	ATCTGAGGTCAGCGAC	CAGCGCATCCGT	1370					
Db	961	ACAAGAATGGCTGTGA	AGGCTCAGCATCACAC	ATCTGAGGTCAGCGAC	CAGCGCATCCGT	1020					
QY	1371	GTGACCAGTGAAGTT	TCTCATTGCATTAA	AGCTGATTAATAAT	GTACATGGGAGAA	1430					
Db	1021	GTGACCAGTGAAGTT	TCTCATTGCATTAA	AGCTGATTAATAAT	GTACATGGGAGAA	1080					
QY	1431	TTTGCAAAAATCAT	TGAAGACCTAAGA	AGGAAGAAAGAA	ACTATTGGAGAAG	1490					
Db	1081	TTTGCAAAAATCAT	TGAAGACCTAAGA	AGGAAGGAAGAA	AGCTATTGGAGAAG	1140					
QY	1491	CTTGTCAGAGCCT	GCACAAGTATAA	CCCTGTTTCATCAT	CCCCACAGTGGCC	1550					
Db	1141	CTTGTCAGAGCCT	GCACAAGTATAA	CCCTGTTTCATCAT	CCCCACAGTGGCC	1200					
QY	1551	TGGGTTCTCATCC	ACACATCCTTA	AAAGTGAAGCTCA	CAGCGTCAATGGC	1610					
Db	1201	TGGGTTCTCATCC	ACACATCCTTA	AAAGTGAAGCTCA	CAGCGTCAATGGC	1260					
QY	1611	CTGGCCTCCTTGA	ATCTCCTTCGG	CTGTCAAGTGTTC	TTTGTGCCATTAT	1670					
Db	1261	CTGGCCTCCTTGA	ATCTCCTTCGG	CTGTCAAGTGTTC	TTTGTGCCATTAT	1320					
QY	1671	CTCACGAATTCCA	AGTCTGCAGTGA	TGAGGTTCAAGA	AGTTTTCCTCC	1730					
Db	1321	CTCACGAATTCCA	AGTCTGCAGTGA	TGAGGTTCAAGA	AGTTTTCCTCC	1380					
QY	1731	GTTTCTATGTCC	AGACATTA	CAAGACCC	CAGCAAAAGCTCT	1790					
Db	1381	GTTTCTATGTCC	AGACATTA	CAAGACCC	CAGCAAAAGCTCT	1440					
QY	1791	TTGTCATGGCA	ACAGACCTGT	CCCGGATCGTCA	ATGGGCAC	1850					
Db	1441	TTGTCATGGCA	ACAGACCTGT	CCCGGATCGTCA	ATGGGCAC	1500					
QY	1851	GGGCATGCTTCT	GAGGGATGAC	CCAGSCCTAG	AGATGCCCTCG	1910					
Db	1501	GGGCATGCTTCT	GAGGGATGAC	CCAGSCCTAG	AGATGCCCTCG	1560					
QY	1911	AACAGCCTGGG	CCCCAGAGTTGC	ACACAAGATCA	ACCTGGTGTCC	1970					
Db	1561	AACAGCCTGGG	CCCCAGAGTTGC	ACACAAGATCA	ACCTGGTGTCC	1620					
QY	1971	GGGTCTCGG	CAACACGCGG	GAGTGGTAAG	AGCAGCCCTGT	2030					
Db	1621	GGGTCTCGG	CAACACGCGG	GAGTGGTAAG	AGCAGCCCTGT	1680					
QY	2031	ATGCATTGCTCG	AGGGCTCGGTGG	GGTGCAGGGA	AGCCCTGGCCTAT	2090					
Db	1681	ATGCATTGCTCG	AGGGCTCGGTGG	GGTGCAGGGA	AGCCCTGGCCTAT	1740					
QY	2091	GCCTGGATCGT	CAGCGGAACAT	CAGGAGAACAT	CCTCATGGG	2150					
Db	1741	GCCTGGATCGT	CAGCGGAACAT	CAGGAGAACAT	CCTCATGGG	1800					
QY	2151	GCCCGATACCT	CCAGGTGCTCC	ACTGCTGCCCTGA	ATCGGGACCTG	2210					
Db	1801	GCCCGATACCT	CCAGGTGCTCC	ACTGCTGCCCTGA	ATCGGGACCTG	1860					
QY	2211	TTTGGAGACAT	GACAGAGATTGG	AGACCGGGGCC	CTCAACCTCT	2270					
Db	1861	TTTGGAGACAT	GACAGAGATTGG	AGACCGGGGCC	CTCAACCTCT	1920					
QY	2271	AGGATCAGCCT	GGCCCCCGCG	CGCTTATTCG	ACCGTCAGATCT	2330					
Db	1921	AGGATCAGCCT	GGCCCCCGCG	CGCTTATTCG	ACCGTCAGATCT	1980					
QY	2331	CTGCTGCTGTGG	ACGCCCCACG	TCGGGAAGCACAT	TTTTGAGGAG	2390					
Db	1981	CTGCTGCTGTGG	ACGCCCCACG	TCGGGAAGCACAT	TTTTGAGGAG	2040					


```
QY      2391 CTCAGGGGGAAGACGGTCGTCCTGGTGACCCACCAGCTGCAGTACTTAGAAATTTGTGGC 2450
      |||||||
Db      2041 CTCAGGGGGAAGACGGTCGTCCTGGTGACCCACCAGCTGCAGTACTTAGAAATTTGTGGC 2100
      |||||||

QY      2451 CAGATCATTTTGTGGAAAAATGGGAAAAATCTGTGAAAAATGGAACCTCACAGTGAGTTAATG 2510
      |||||||
Db      2101 CAGATCATTTTGTGGAAAAATGGGAAAAATCTGTGAAAAATGGAACCTCACAGTGAGTTAATG 2160
      |||||||

QY      2511 CAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAGCCACTTCGG 2568
      |||||||
Db      2161 CAGAAAAAGGGGAAATATGCCCAACTTATCCAGAAGATGCACAAGGAGCCACTTCGG 2218
      |||||||
```

Search completed: July 24, 2003, 09:31:52
Job time : 12076 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2003, 05:20:50 ; Search time 6169 Seconds
(without alignments)
12764.217 Million cell updates/sec

Title: US-10-087-782A-1
Perfect score: 4862
Sequence: 1 actgggataaagcaaga.....tataaaactaaggaaaactc 4862

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*			
1:	em_estba:*		
2:	em_esthum:*		
3:	em_estin:*		
4:	em_estmu:*		
5:	em_estov:*		
6:	em_estpl:*		
7:	em_estro:*		
8:	em_htc:*		
9:	gb_est1:*		
10:	gb_est2:*		
11:	gb_htc:*		
12:	gb_est3:*		
13:	gb_est4:*		
14:	gb_est5:*		
15:	em_estfun:*		
16:	em_estom:*		
17:	gb_gss:*		
18:	em_gss_hum:*		
19:	em_gss_inv:*		
20:	em_gss_pln:*		
21:	em_gss_vrt:*		
22:	em_gss_fun:*		
23:	em_gss_mam:*		
24:	em_gss_mus:*		
25:	em_gss_other:*		
26:	em_gss_pro:*		
27:	em_gss_rod:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	430.8	8.9	559	BI462847	BI462847 603207826
2	381.2	7.8	398	BE089591	BE089591 QV0-BT070
3	349.2	7.2	367	AI401832	AI401832 th24b12.x
4	324.2	6.7	667	BJ499217	BJ499217 BJ499217
5	323	6.6	325	BF447217	BF447217 7p46d02.x
6	310.2	6.4	326	AI676121	AI676121 wc05f11.x

7	289	5.9	801	9	AJ393794	AJ393794
8	279.8	5.8	550	12	BF190023	BF190023 236002 MA
9	273.2	5.6	761	9	AU137090	AU137090 AU137090
10	272.2	5.6	935	14	BQ898022	BQ898022 AGENCOURT
11	258.6	5.3	900	14	BQ891357	BQ891357 AGENCOURT
12	253.8	5.2	888	12	BE869517	BE869517 601446187
13	252.8	5.2	578	14	BQ267358	BQ267358 ik02g05.y
14	249	5.1	567	13	BM512586	BM512586 ij70q03.y
15	230.4	4.7	888	14	BQ687464	BQ687464 AGENCOURT
16	229.8	4.7	674	14	BP007828	BP007828 BP007828
17	229.2	4.7	1006	14	BM804774	BM804774 AGENCOURT
18	228.6	4.7	896	14	BQ434553	BQ434553 AGENCOURT
19	225.8	4.6	584	13	BI775730	BI775730 468418 MA
20	222.4	4.6	520	12	BE754240	BE754240 207698 MA
21	222.2	4.6	572	13	BI843062	BI843062 fr02f04.y
22	219.8	4.5	1449	11	AY107851	AY107851 Zea mays
23	217.8	4.5	606	14	BM978938	BM978938 UI-CF-DU1
24	215.6	4.4	898	14	BQ690370	BQ690370 AGENCOURT
25	214.6	4.4	1456	11	AY106986	AY106986 Zea mays
26	214.6	4.4	1977	11	U66674	U66674 Homo sapien
27	213.8	4.4	594	9	AJ397153	AJ397153 AJ397153
28	213.2	4.4	675	14	BM936981	BM936981 UI-M-CD0-
29	212.2	4.4	1124	9	AU252331	AU252331 AU252331
30	210	4.3	600	13	BI990628	BI990628 4073-57 M
31	207.6	4.3	237	10	AW372855	AW372855 RC5-BT037
32	207.6	4.3	237	10	AW372856	AW372856 RC5-BT037
33	207.6	4.3	250	10	AW372862	AW372862 RC5-BT037
34	204.8	4.2	1914	11	U66684	U66684 Homo sapien
35	204.6	4.2	690	9	AL694493	AL694493 AL694493
36	204.4	4.2	983	11	AK006128	AK006128 Mus muscu
37	204.4	4.2	998	14	BQ686840	BQ686840 AGENCOURT
38	202.6	4.2	696	12	BG328614	BG328614 602427877
39	201.2	4.1	627	14	BQ814444	BQ814444 1030043E0
40	200.2	4.1	618	10	AV661510	AV661510 AV661510
41	199.8	4.1	800	12	BG205053	BG205053 RST24472
42	199.2	4.1	747	13	BM408634	BM408634 EST582961
43	199	4.1	740	14	BQ769313	BQ769313 UI-M-FIO-
44	198	4.1	413	12	BE860035	BE860035 UI-M-AO0-
45	196.8	4.0	496	9	AA570953	AA570953 vl88e11.r

ALIGNMENTS

RESULT 1	BI462847	BI462847	559 bp	mRNA	linear	EST 21-AUG-2001
BI462847	603207826	Fl NIH_MGC_97	Homo sapiens	cdna	clone IMAGE:5273407 5',	
LOCUS	BI462847	mRNA sequence.				
DEFINITION	BI462847	EST.				
ACCESSION	BI462847.1	GI:152535503				
VERSION	BI462847.1	GI:152535503				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	1 (bases 1 to 559)					
TITLE	NIH-MGC http://mgc.nci.nih.gov/.					
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)					
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cdna Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM11690 row: c column: 08 High quality sequence stop: 554. Location/Qualifiers					
FEATURES						


```
source      1. .559
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5273407"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 130 a 132 c 137 g 160 t
ORIGIN

Query Match      8.9%; Score 430.8; DB 13; Length 559;
Best Local Similarity 99.3%; Pred. No. 1.1e-113;
Matches 443; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY      894 ATATTGATTATACCAAGATCCTGGAAATATTCAAGAAGACGAGTGGGGAATGTTGTCCAT 953
Db      |||||||
QY      954 GGAGTGGGACTCTGCTTGCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTCTCC 1013
Db      |||||||
QY      171 GGAGTGGGACTCTGCTTGCCTTTTCTCTCCGAATGTGTGAAGTCTCTGAGTTTCTCC 230
Db      |||||||
QY      1014 TCCAGTTGGATCATCAA-CCAAAGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCCTT 1072
Db      |||||||
QY      231 TCCAGTTGGATCATCAAGCCAATGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCTCCTT 290
Db      |||||||
QY      1073 TGCCTTTGAGAAGCTCATCCAATTTAAGTCTGTGAATACACATCACCTCAGGAGAGGCCAT 1132
Db      |||||||
QY      291 TGCCTTTGAGAAGCTCATCCAATTTAAGTCTGTGAATACACATCACCTCAGGAGAGGCCAT 350
Db      |||||||
QY      1133 CAGCTTCTTCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTAGT 1192
Db      |||||||
QY      351 CAGCTTCTTCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTAGT 410
Db      |||||||
QY      1193 ACTGATCACCTGGCAATCGCTGGTGCATCTGCAGCAATTTCTTCCTACTTCAATTATGGGATA 1252
Db      |||||||
QY      411 ACTGATCACCTGGCAATCGCTGGTGCATCTGCAGCAATTTCTTCCTACTTCAATTATGGGATA 470
Db      |||||||
QY      1253 CACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCACCTGGCGGTATTCATGAC 1312
Db      |||||||
QY      471 CACTGCATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCACCTGGCGGTATTCATGAC 530
Db      |||||||
QY      1313 AAGAATGGCTGTGAAGGCTCAGCATC 1338
Db      |||||||
QY      531 AAGAATGGCTGTGAAGGCTCAGCATC 556

RESULT 2
BE089591
LOCUS      QV0-BT0703-280400-211-e08 BT0703 Homo sapiens cDNA, linear EST 12-JUN-2000
DEFINITION BE089591
ACCESSION  BE089591
VERSION    BE089591.1 GI:8480009
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 398)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brownstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
```

```
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
COMMENT    Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-st2-QV0-BT0703-280
400-211-e08&t3=2000-04-28&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 398.
FEATURES   Location/Qualifiers
            1. .398
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="BT0703"
                /dev_stage="Adult"
                /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                SmaI; A mini-library was made by cloning products derived
                from ORESTES PCR (U.S. Letters Patent application No. 196
                ,716 - Ludwig Institute for Cancer Research) profiles
                into the pUC 18 vector. Reverse transcription of tissue
                mRNA and cDNA amplification were performed under low
                stringency conditions."
BASE COUNT 93 a 103 c 86 g 116 t
ORIGIN

Query Match      7.8%; Score 381.2; DB 10; Length 398;
Best Local Similarity 99.0%; Pred. No. 2.5e-99;
Matches 394; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY      1018 GTTGGATCATCAACCAACGC-ACAGCCATCAGGTTCCGAGCAGCTGTTTCCCTTGCC 1076
Db      | |||||||
QY      1  GGTGGATCATCAACCAACGCACAGCCATCAGGTTCCGAGCAGCTGTTTCCCTTGCC 60
Db      |||||||
QY      1077 TTTGAGAAGCTCATCCAATTTAAGTCTGTGAATACACATCACCTCAGGAGAGGCCATCAGC 1136
Db      |||||||
QY      61  TTTGAGAAGCTCATCCAATTTAAGTCTGTGAATACACATCACCTCAGGAGAGGCCATCAGC 120
Db      |||||||
QY      1137 TTCTTCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTAGTACTG 1196
Db      |||||||
QY      121  TTCTTCACCGGTGATGTAAACTACCTGTTTGAAGGGGTGTGCTATGGACCCCTAGTACTG 180
Db      |||||||
QY      1197 ATCACCTGGCATCGCTGGTGCATCTGCAGCATTTCTTCCTACTTCAATTATGGATACACT 1256
Db      |||||||
QY      181  ATCACCTGGCATCGCTGGTGCATCTGCAGCATTTCTTCCTACTTCAATTATGGATACACT 240
Db      |||||||
QY      1257 GCATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCACCTGGCGGTATTCATGACAAGA 1316
Db      |||||||
QY      241  GCATTTATTGCCATCTTATGCTATCTCCTGGTTTCCCACCTGGCGGTATTCATGACAAGA 300
Db      |||||||
QY      1317 ATGGCTGTGAAGGCTCAGCATCACACATCTGAGTCAAGCCAGCGCATCGGTGTGACC 1376
Db      |||||||
QY      301  ATGGCTGTGAAGGCTCGGCATCACACATCTGAGTCAAGCCAGCGCATCGGTGTGACC 360
Db      |||||||
QY      1377 AGTGAAGTTCTCACTGCATTAAGCTGATTAAGTGA 1414
Db      |||||||
QY      361  AGTGAAGTTCTCACTGCATTAAGCTGATTAAGTGA 398
Db      |||||||

RESULT 3
AI401832/c
LOCUS      AI401832
DEFINITION th24b12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2119199 3',
mRNA sequence.
mRNA
```

ACCESSION AI401832
VERSION AI401832.1 GI:4244919
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 367)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 979 Std Error: 0.00
Seq primer: -40Up from Gibco.
FEATURES
source
1..367
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2119199"
/clone_lib="NCI_CGAP_Pr28"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI_CGAP_Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 99 a 96 c 68 g 104 t
ORIGIN
Query Match 7.2%; Score 349.2; DB 9; Length 367;
Best Local Similarity 99.2%; Pred. No. 5.2e-90;
Matches 351; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4509 GAGACTTCATGGAGGCTGGCAGCTCAGAGGTTACACAGGTGCAGCTTCGAGGCC 4568
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 356 GAGACTTATAGGAGGCTGGCAGCTGAGCTCAGAGGTTACACAGGTGCAGCTTCGAGGCC 297
QY 4569 CACAGTCTGCGACCTTCTTGTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAATGT 4628
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 296 CACAGTCTGCGACCTTCTTGTTGGAGATGAGAACTTCTCCTGGAAGCAGGGGTAATGT 237
QY 4629 AGGGGGGTGGGGATTGCTGGATGGAAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGA 4688
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 236 AGGGGGGTGGGGATTGCTGGATGGAAACCCCTGGAATAGGCTACTTGATGGCTCTCAAGA 177
QY 4689 CCTTAGAACCCCGAACCATCTAAGACATGGGATTTCAGTGATCATGTGGTTCTCCTTTTA 4748
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 176 CCTTAGAACCCCGAACCATCTAAGACATGGGATTTCAGTGATCATGTGGTTCTCCTTTTA 117
QY 4749 ACTTACATGCTGAATAATTTTATAAAGGTAAAGCTTATAGTTTCTGTATCTGTGTTA 4808
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 116 ACTTACATGCTGAATAATTTTATAAAGGTAAAGCTTATAGTTTCTGTATCTGTGTTA 57
QY 4809 GAAGTGTGCAAATGCTGTACTGACTTTTGTAATAATATAAACTAAGGAAACTC 4862
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 56 GAAGTGTGCAAATGCTGTACTGACTTTTGTAATAATATAAACTAAGGAAACTC 3

RESULT 4
BJ499217 667 bp mRNA linear EST 08-AUG-2002
LOCUS BJ499217 MF01FSA cDNA Oryzias latipes cDNA clone MF01FSA018J02 5',
DEFINITION mRNA sequence.
ACCESSION BJ499217
VERSION BJ499217.1 GI:22151179
KEYWORDS EST.
SOURCE Japanese medaka.
ORGANISM Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE 1 (bases 1 to 667)
AUTHORS Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLE Medaka EST Project in Takeda's lab
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
1..667
/organism="Oryzias latipes"
/strain="d-rR"
/db_xref="taxon:8090"
/clone="MF01FSA018J02"
/clone_lib="MF01FSA cDNA"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
BASE COUNT 174 a 185 c 176 g 128 t 4 others
ORIGIN
Query Match 6.7%; Score 324.2; DB 13; Length 667;
Best Local Similarity 68.7%; Pred. No. 1.7e-82;
Matches 443; Conservative 0; Mismatches 202; Indels 0; Gaps 0;
QY 3843 ATCCGGCGCCACGAAGTGGTGGCATCGTGGGAAGGACGGGCTCTGGGAAGTCCTCTTG 3902
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 ATTCGGCACGAGGAGAAACTTGGCATTGTGGGGCGAACAGGCTCAGGAAAGTCCTCTTTA 60
QY 3903 GGCATGGCTCTCTTCCGCCCTGGTGGAGCCCATGGCAGGCCGGAATTCTCATTTGACGGCGTG 3962
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 GCGTCGCTCTGTTCAGAAATGGTGGAGCCAGCAGCAGGAACCATAGTATCGATGGCGTG 120
QY 3963 GACATTTGCAGCATCGGCCCTGGAGGACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAGAT 4022
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 GACATCTCAGGTATCGGCCCTGAAGGACCTGCGCAGCAAACTGTCCATCATCCCCGACGAC 180
QY 4023 CCAGTGCTGCTCTCTCAGGAACCATCAGATTCAAACCTAGATCCCTTTGACCGTCACACTGAC 4082
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 181 CCTGTGCTGTTATTGGCACACTCAGATACAACCTGGACCCCTTCGATAACTACTCTGAC 240
QY 4083 CACCAGATCTGGGATGCCCTTGGAGAGGACATTCCTGACCAGGCCATCTCAAAGTTCCCC 4142
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 GAGGAGATCTGGAAGGCTCTGGAGAAGACCTACATGAAGGACTCGATCTCCATGCTGGAG 300
QY 4143 AAAAAGCTGCATACAGATGTGGTGGAAAACGGTGGAAACTTCCTGACCAAGGCCATCTCAAAGTTCCCC 4202
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 301 GATAAACTGCAGGCCCAAAGTGTGGAGAACCGAGAGAACTTCTCTGTAGCGGAGAGGCAG 360
QY 4203 CTGCTCTGCATTGCCAGGGCTGTGCTTCGAACTCCAAGATCATCTTATCGATGAAGCC 4262
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 361 CTGATCTGCATGGCCAGAGCACTTCTGCGCAACTCCNAGATCATCTTCTGGATGAGGCA 420
QY 4263 ACAGCCTCCATTGACATGGAGACAGACACCCCTGATCCAGCGCACAATCCGTGAAGCCTTC 4322
||||||| || ||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||

Db 421 ACAGCCTCAATCGACGAGAAACAGATACTCTGATCCAAAACACCATCAAGGAGGCCTTC 480

QY 4323 CAGGGCTGCACCGTGCTCATTTGCCACCCTGTTCACCACTGTGCTGAACCTGTGACCCAC 4382

Db 481 CACCACTGCACCATGCTCACCATCGCGCACCCGANCANCAACACGGTCTCATGCACTCGGATCGG 540

QY 4383 ATCCTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCCGGAGGTACTGCGGAAG 4442

Db 541 ATCCTGGTCATGGACAGAGCGGAGGCAGCAGAACTGGACCGTCCAGAGGTGCTGAGGCCAA 600

QY 4443 AAGCCTGGGTCATTTGTCGACGCCCTCATGGCCACAGCCACTTCT 4487

Db 601 AGACCCAACTCCCTGTTCTCCTCACTCCTCGCTGCAGCTNANACT 645

RESULT 5

BF447217/c

LOCUS 7p46d02.x1 NCI_CGAP_Pr28 325 bp mRNA linear EST 01-DEC-2000

DEFINITION mRNA sequence.

ACCESSION BF447217

VERSION BF447217.1 GI:11512355

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 325)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco.

FEATURES

1. .325

Location/Qualifiers

1. .325

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3648818"

/clone_lib="NCI_CGAP_Pr28"

/sex="male"

/dev_stage="adult"

/lab_host="DH10B"

/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 93 a 81 c 58 g 93 t

ORIGIN

Query Match 6.6%; Score 323; DB 12; Length 325;

Best Local Similarity 100.0%; Pred. No. 2.1e-82;

Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4540 GAGGTTACACAGGTGCAGCTTCGAGGCCACAGTCTCGGACCTTCTGTTGGAGATGA 4599

Db 325 GAGGTTACACACAGGTGCAGCTTCGAGGCCACAGTCTCGGACCTTCTGTTGGAGATGA 266

QY 4600 GAACTTCTCCTGGAAGCAGGGGTAAATGTAGGGGGGTGGGGATTGCTGCGATGGAAACCC 4659

Db 265 GAACTTCTCCTGGAAGCAGGGTAAATGTAGGGGGTGGGATTGCTGGTGAACCC 206

QY 4660 TGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCGAGAACCATCTAAGACATGG 4719

Db 205 TGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCGAGAACCATCTAAGACATGG 146

QY 4720 GATTCAGTGATCATGTGGTTCTCCTTTTAACCTTACATGCTGAATAATTTATAATAAGGT 4779

Db 145 GATTCAGTGATCATGTGGTTCTCCTTTTAACCTTACATGCTGAATAATTTATAATAAGGT 86

QY 4780 AAAAGCTTATAGTTTCTGATCTGTGTAGAGTGTTCGAAATGCTGTACTGACTTTGTA 4839

Db 85 AAAAGCTTATAGTTTCTGATCTGTGTAGAGTGTTCGAAATGCTGTACTGACTTTGTA 26

QY 4840 AAATATAAACTAAGGAAAACTC 4862

Db 25 AAATATAAACTAAGGAAAACTC 3

RESULT 6

AI676121/c

LOCUS wc05f11.x1 NCI_CGAP_Pr28 326 bp mRNA linear EST 17-DEC-1999

DEFINITION mRNA sequence.

ACCESSION AI676121

VERSION AI676121.1 GI:4876601

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 326)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert length: 295 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES

Location/Qualifiers

1. .326

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2314317"

/clone_lib="NCI_CGAP_Pr28"

/sex="male"

/dev_stage="adult"

/lab_host="DH10B"

/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 94 a 81 c 61 g 90 t

ORIGIN

Query Match 6.4%; Score 310.2; DB 9; Length 326;

Best Local Similarity 99.0%; Pred. No. 1.2e-78;

Matches 312; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	4540	GAGGTTACACACAGGTGCAGCTTCGAGGCCACAGCTCTCGACCTTCTTGTGGAGATGA	4599
Db	326	GAGGTTACACACAGGTGCAGCTTCGAGGCCACAGCTCTCGACCTTCTTGTGGAGATGA	267
QY	4600	GAACTTCTCCTGGAAGCAGGGTAAATGTAGGGGGGTGGGATTGCTGGATGGAACCC	4659
Db	266	GAACTTCTCCTGGAAGCAGGGTAAATGTAGGGGGGTGGGATTGCTGGATGGAACCC	207
QY	4660	TGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCGAACCATCTAAGACATGG	4719
Db	206	TGGAATAGGCTACTTGATGGCTCTCAAGACCTTAGAACCCCGAACCATCTAAGACATGG	147
QY	4720	GATTCAGTGATCATGTGGTTCTCCTTTTAACCTTACATGCTGAATAATTTTATAATAAGGT	4779
Db	146	GATTCAGTGATCATGTGGTTCTCCTTTTAACCTTACATGCTGAATAATTTTATAATAAGGT	87
QY	4780	AAAAGCTTATAGTTTCTGATCTGTGTTAGAAAGTGTGCAAAATGCTGACTGACTTTGTA	4839
Db	86	AAAAGCTTATAGTTTCTGATCTGTGTTAGAAAGTGTGCAAAATGCTGACTGACTTTGTA	27
QY	4840	AAATATAAAACTAAG	4854
Db	26	AAATATAAAACTAAG	12
RESULT 7			
AJ393794			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
COMMENT			
FEATURES			
source			
1..801			
/organism="Gallus gallus"			
/strain="CB"			
/db_xref="taxon:9031"			
/clone="16ol5rl"			
/clone_lib="dkfz426"			
/tissue_type="Bursa of Fabricius"			
/cell_type="bursal lymphocyte"			
/dev_stage="2-3 weeks old"			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
445; Conservative			
0; Mismatches			
260; Indels			
0; Gaps			
0;			
QY	3735	ACAAAGTTGTCGCCAGGGTGGCCACAGCATGGGGAAATCATATTTACGATTATCACATG	3794
Db	4	AAAACTCCCCCTCTGGACTGGCCACAGGAAGTGAAGTTGTTTTGAAAAATGCAGAGATG	63
QY	3795	AAATACAGAGACACACACCCACCGTGTTCACGGCATCAACCTGACCATCCGCGCCAC	3854

Db	64	CGGTACCGAGAGAAATCTCCCTCTGGTACTTTAAAAAAGTGTCTCTTACCATCAAAACCAAAG	123
QY	3855	GAAGTGGTGGCATFCGTGGGAAGGACGGGCTCTCGGGAAGTCTCTTGGGATGGCTCTC	3914
Db	124	GAAAAGATTGGCATTTGTGGGAAGGACAGGCTCAGGGAAGTCTTCCCTTGAATGGCACTC	183
QY	3915	TTCCGCCCTGGTGAGCCCCATGGCAGGCGCGGATTTCTCATTTGACGGGTGGACATTTGCAGC	3974
Db	184	TTTTCGTTTGGTTGAGCTATCTGGAGGCTGCAATTAATAATTGATGGAGTGAATAAACGAC	243
QY	3975	ATCGGCCCTGGAGGACTTGCGGTCCAAGCTCTCAGTGATCCCTCAAGATCCAGTGTCTC	4034
Db	244	ATTGGTTTGGCAGATCTTCGCAGCAAACCTCTCAATAATCCCTCAGGAACCTGTGCTGTT	303
QY	4035	TCAGGAACCATCAGATTTCAACCTAGATCCCTTTGACCCGTACACACTGACCAGCAGATCTGG	4094
Db	304	AGTGGCACCCGTAAGATCAAACTTGGATCCCTTCAATCAATACAGTGAGGAGCAAATCTGG	363
QY	4095	GATGCCCTTGGAGAGGACATTCCTGACCAAGGCCATCTCAAAAGTTCGCCAAAAGCTGCAT	4154
Db	364	GTTGCCCTTGGAAAGGACTCACATGAAGGAGTGTGTTGCCAGCTGCCTATGAAACTTCGAC	423
QY	4155	ACAGATGTGGTGGAAAAACGGTGGAAACTTCTCTGTGGGGGAGAGGCAGCTGCTGCATTT	4214
Db	424	TCAGAAGTAATGAAAAATGGGAAAACTTTTCAGTTGGAGAGAGGCAGTTGCTGTGCATA	483
QY	4215	GCCAGGGCTGTGCTTCGCAACTCCAAGATCATCTTATCGATGAAGCCACAGCCTCCATT	4274
Db	484	GCTCGAGCTCTGCTGCGTCGCTGCAAGATTTTGATACTGGATGAAGCAACAGCTGCTATG	543
QY	4275	GACATGGAGACACACACCCTGATCCAGCGCACAAATCCGTGAAGCCTTCCAGGGCTGCACC	4334
Db	544	GACACAGAGACTGACTTGCTGATTCAGGAGACCATCAGAGAGGCATTTGCAGACTGCACT	603
QY	4335	GTGCTCGTCATTGCCCCACCGTGTCAACCACITGCTGAACTGTGACCACATCCTGGTTATG	4394
Db	604	ATGTTAACAAATTGCTCATCGCCTGCATACITGTA	663
QY	4395	GGCAATGGGAAGGTGGTAGAATTTGATCGGCGGAGGAGGTACTGCGG	4439
Db	664	ACACAAGGACAGGTAGTGGAAATTCGACACCGCGTCTGCCCTTCTG	708
RESULT 8			
BF190023			
LOCUS	BF190023	550 bp	mrna linear EST 02-NOV-2000
DEFINITION	236002 MARC 2PIG Sus scrofa cdna 5', mRNA sequence.		
ACCESSION	BF190023		
VERSION	BF190023.1 GI:11073392		
KEYWORDS	EST.		
SOURCE	pig.		
ORGANISM	Sus scrofa		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.		
AUTHORS	1 (bases 1 to 550)		
	Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.		
TITLE	Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine		
JOURNAL	unpublished (2000)		
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options. PCR Primers FORWARD: AGGAAACAGCTATGACCAT BACKWARD: GTTTTCCCAGTCACGACG		

Plate: 60 row: E column: 3
Seq primer: ATTTAGGTGACACTATAG.
FEATURES
source
1. .550
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2PIC"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 134 a 135 c 153 g 128 t
ORIGIN
Query Match 5.8%; Score 279.8; DB 12; Length 550;
Best Local Similarity 79.2%; Pred. No. 1.4e-69;
Matches 332; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
QY 510 GGGAGGCGAGCTGTCCACCGTGGGGAAGTATGATGCTGCCTTGAGAAACCATGATTCCTC 569
Db 132 GGGAGCCTGAGGCGCCAAGGATGGGAAGTATGATGCTGCCTGGAGAACCATGATTCCTC 191
QY 570 TTCCGTCCCAAGCCGAGGTTTCCTGCCCCCAGCCCCCTGGACAATGCTGGCCTGTTCTCC 629
Db 192 TTCCGCCCCGAAGCCAAAGATTCTCTGCCCCCAGCCCCCTGGATGATGCGCGCTTGTCTCC 251
QY 630 TACCTCACCGTGTTCATGGCTCACCCCGCTCATGATCCAAAGCTTACGGAGTCGCTTAGAT 689
Db 252 TACCTCACCTGTGCTGGCTCACCCCATTCATGATCCGAGGTTTACGGAAACGCTTAAAT 311
QY 690 GAGAACACCATCCCTCCACTGTCACTGCCATGATGCCCTCAGACAAAAATGTCCAAAGGCTT 749
Db 312 AAGCACATGATCCCCCAGCTGTCACTGATGATGCTTTCAGCCCCAAAATGCCAAAGGCTT 371
QY 750 CACCGCCTTTGGGAAGAAGTCTCAAGGCGAGGGATTGAAAAAGCTTCAGTGTCTCTG 809
Db 372 CGTCTCCTCTGGGAAGAAGTCTCAAGGCATGGGATTGACAAAGCGTCGGTGTTCGA 431
QY 810 GTGATGCTGAGGTTCCAGAGAACAAAGTTGATTTTCGATGCACCTTCTGGGCATCTGCTTC 869
Db 432 GTGATGCTGAAATTCCAAAGAACAAAGGTTATTTTGTATGATGGTCTCGAGCTGTTCTTC 491
QY 870 TGCATTGCCAGTGTACTCGGGCCCAATATTGATTATACCAAAGATCCTTGGAAATATCAGA 928
Db 492 TCTGCCACACGCGTGCTATGGCCCAATGCTGGTTATACCAAAGATCCTTGGAAATATGCAGA 550
RESULT 9
AUI37090
LOCUS
DEFINITION
AUI37090 PLACEL1 Homo sapiens cDNA clone PLACEI005753 5', mRNA
sequence.
ACCESSION
VERSION
AUI37090.1 GI:10997629
KEYWORDS
SOURCE
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 761)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE
HRI human cDNA project
JOURNAL
Unpublished (2000)
COMMENT
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Location/Qualifiers
1. .761
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACEI005753"
/clone_lib="PLACEI"
/tissue_type="placenta"
/note="Vector: pMEI8SFL3"
BASE COUNT 200 a 185 c 204 g 165 t 7 others
ORIGIN
Query Match 5.6%; Score 273.2; DB 9; Length 761;
Best Local Similarity 63.6%; Pred. No. 1.5e-67;
Matches 430; Conservative 0; Mismatches 245; Indels 1; Gaps 1;
QY 3745 CCCAGSGGTGCCACAGCATGGGGAATCATATTTTCAGGATTATACATGAAATACAGAG 3804
Db 17 CCCCTGACTGGCCCCCAGGAGGAGGTGACCTTTTGAGAACGCAGAGATGAGGTACCGAG 76
QY 3805 ACAACACACCCACCCTGCTTCACGGCATCAACCTGACCATCCGCGCCACGAAGTGGTGG 3864
Db 77 AAACCTCCCTCTCGTCTCTAAAGAAAGTATCCTTCACGATCAAACCTTAAGAGAAGATTG 136
QY 3865 GCATCGTGGGAAGGACGGGCTCTGGGAAGTCTCTCTTGGGCATGGCTCTCTCCGCCTGG 3924
Db 137 GCATTGTGGGCGGACAGGATCAGGAAAGTCTCTCGCTGGGGATGGCCCTCTTCCGTCTGG 196
QY 3925 TGGAGCCCATGGCAGGCCGGATTCTCATTTGACGGCGTGGACATTTGCAGCATCGGCCTGG 3984
Db 197 TGGAGTTATCTGGAGGCTGCATCAAGATTGATGGAGTGAGAATCAGTGATATTGGCCTTG 256
QY 3985 AGGACTTTCGGTCCCAAGCTCTCACTGATCCCTCAAGATCCAGTGTGCTCTCAGGAACCA 4044
Db 257 CCGACCTCCGNAGCAAACTCTCTATCATTCCTCAAGAGCCGGTGTGTTCAGTGGCAGTG 316
QY 4045 TCAGATTCAACCTTAGATCCCTTTTGACCGTCACACTGACCAGCAGACTTGGGATGCCCTTGG 4104
Db 317 TCAGATCAAAATTGGACCCCTTCACACCAGTACACTGAAGACCAGATTTGGGATGCCCTGG 376
QY 4105 AGAGGACATTCCTGACCAAGGCCATCTCAAAGTTCCCCAAAAAGCTGCATACAGATGTGG 4164
Db 377 AGAGGACACACATGAAAGAAATGTATTGCTCAGCTACCTCTGAAACTTGAATCTGAAGTGA 436
QY 4165 TGGAAAACGGTGGAACCTTCTCTGTGGGGGAGAGGCGAGCTGCTCTGCATTTGCCAGGGCTG 4224
Db 437 TGGAGAATGGGGATAAATCTCTCAGTGGGGGAACGGCAGCTCTTGTGCATAGCTAGAGCCC 496
QY 4225 TGCTTCGCAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCTCCATTGACATGGAGA 4284
Db 497 TGCTCCGCCACTGTAAGATTCTGATTTTAGATGAAGCCACAGCTGCCATGGACACAGAGA 556
QY 4285 CAGACACCCCTGATCCAGCGCACAAATCCGTGAAGCCTTCCAGGGCTGCACCGTGCCTCTCA 4344
Db 557 CAGACTTATTGATTCAAGAGACCACATCCGAGAAGCATTTGCAAACTGTACCAAGCTGACCA 616
QY 4345 TTGCCACCCTGTCCACCCTGTGCTGAACTGTGA-CCACATCCTTGGTTATGGGCAATGGG 4403
Db 617 ATGCCCATCGNCTGCACACCGGGTCTAAGCTCCGATAGGGATTATGGGGCTGGGCCAAGGA 676
QY 4404 AAGGTGGTAGAATTTG 4419
Db 677 CAAGTGGTGAGTTTG 692
RESULT 10
BQ898022
LOCUS
DEFINITION
BQ898022 935 bp mRNA linear EST 16-AUG-2002
AGENCOURT_8122242 Lupski_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6179508 5', mRNA sequence.

Db 63 GGCCTCTCCTCATCCTCTTCTCACTTCTCCACATGTCTCCAGGGTCCGTGATTCGTGAGC 122

QY 3310 TCAAGAGACTGGAGAACTATAGCCGGTCTCCTTTATTCTCCACATCCTCAATTCTCTGC 3369

Db 123 TAAAGCGTTGGACAATATCACGCGAGTCTCCTTTCTCTCCACATCACGCTCTAGCATTC 182

QY 3370 AAGGCCTGAGCTCCATCCATGTCTATGGA AAAA ACTGAAGACTTTCATCAGCCAGTTTAAGA 3429

Db 183 AGGCGCTGGCCACCATCCATGCCCTACAACAAAGGCAGGAGTTTTACACAGGTATCAGG 242

QY 3430 GGCTGACTGATGCCAGAGAAATAACTACCTGCTGTTGTTTCTATCTTCCACACGATGGATGG 3489

Db 243 AGCTCCTGGATGACAACACGAGCTCCCTTTTCTCCTGTTCACTGTGCAATGAGGTGGCTGG 302

QY 3490 CATTGAGGCTGGAGATCATGACCAACCTTTGTGACCTTGGCTGTTGCTGCTGCTGCTGCTT 3549

Db 303 CAGTGGCTGGACCTCATCAGCATTTGCCCTGATTACCACCACTGGCCCTGATGATTGTTTC 362

QY 3550 TTGGCATTTCTCCACCCCTACTCTTTAAAGTCATGGCTGTCAACATCGTGTGCAGC 3609

Db 363 TCATGCATGGACAGATCCCTTCAGCCTATGCGGGGCTTGCCATCTCCTACGCTGTGCAGT 422

QY 3610 TGGCGTCCAGCTTCCAGGCCACTGCCCGGATTTGGCTTGGAGACAGAGGCACAGTTCACGG 3669

Db 423 TAACTGGACTATTCCAGTTTCACCGTCAGACTGGCATCGGAGACAGAGCACGGTTCACATT 482

QY 3670 CTGTAGAGAGGATACTGCAGTACATGAAGATGTGTCTCGGAAGTCCTTTACACATGG 3729

Db 483 CCGTGGAGAGGATCAACCACTATATCAAGACTCTCTCTTTGGAAGCACCTGCCAGAAATCA 542

QY 3730 AAGGCACAAGTTGTCCCCAGGGTGGCCACAGCATGGGGAATCATATTTCAGGATTATC 3789

Db 543 AGAACAAAGGCTCCTCCCCATGACTGGCCCCAGGAGGAGAAGTAACCTTTGAGAATGCAG 602

QY 3790 ACATGAAATACAGAGACAAACACACCCCGTGTTCACGGCATCAACCTGACCATCCGCG 3849

Db 603 AAATGAGATACCGGGAAAATCTCCCTCTGTCTCCTTAAGAAAGTGTCCTTCACCATCAAGC 662

QY 3850 GCCACGAAGTGTGGGCATCGTGGGAAGCAGCGGCTCTGGGAAGTCCCTTGGGCATGG 3909

Db 663 CCAAGGAAAAGATAGGCATTGTGGACGAACAGGGTCAGGGAAGTCTCTTTGGGGATGG 722

QY 3910 CTCTCTCCGCTGGTGGAGCCCATGGCAGCCCGGATTCTCATTTGACGGCGTGGACATTT 3969

Db 723 CCCTCTCCGCTGGTGGAGCTATCTGGAGGCTGCATCAAGATTGTGGAATAAGAATCA 782

QY 3970 GCAGCATCGGCTGGAGGACTTGGCGGTCCAAGCTCTCAGTGATCCCTCAAGATCCAGTGC 4029

Db 783 GTGACATCGGNTGGNCGACCCCTCGAAACCAACTGGCCATCATTTCTCAGGAGCCAGTGC 842

QY 4030 TGCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTTTGACCGTCCACACTG 4080

Db 843 TGTTCACTGGCACTGTACAGATCAAACTGGACCCCTTTTTCACCACTACACGG 893

RESULT 12

BE869517

LOCUS BE869517 888 bp mRNA linear EST 20-OCT-2000

DEFINITION 601446187F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3850196 5', mRNA sequence.

ACCESSION BE869517

VERSION BE869517.1 GI:10318293

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 888)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC).

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM9568 row: n column: 21

High quality sequence stop: 620.

FEATURES

source

Location/Qualifiers

1..888

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3850196"

/clone_lib="NIH_MGC_65"

/tissue_type="adenocarcinoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: Sali; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.8 kb. Library constructed by Life Technologies."

BASE COUNT 192 a 261 c 269 g 165 t 1 others

ORIGIN

Query Match 5.2%; Score 253.8; DB 12; Length 888;

Best Local Similarity 81.9%; Pred. No. 7.9e-62;

Matches 367; Conservative 0; Mismatches 68; Indels 13; Gaps 6;

QY 332 GAATTGGCAGGAACCTGAAAAATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGG 391

Db 361 GAATTGGCAGGAACCTGAAAAATGACTAGGAAGAGGACATACTGGGTGCCCAACTCTTCTGG 420

QY 392 TGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAA 451

Db 421 TGGCCTCGTGAATCGTGGCATCGACATAGGCGATGACATGGTTTCAGGACTTATTTATAA 480

QY 452 AACCTATACTCTCC-AAGATGGCCCCCTGGAGTCAGCAAGAGAGAAATCCTGAGGCTCCAG 510

Db 481 AACCTATACTCTCCAAAGATGGCCCCCTGGAGTCAGCAAGAGAGACATCCTGAGGCTCCAG 540

QY 511 GGAGGGCAGCTGTCCCACCGTGGGGGAAGTA--TGATGCTGCCCTTGAGAACCATGATTCC 568

Db 541 GGAGGGCAGCTGTCCCACCGTGGCGGAAGTAATGATGCCTGCCCTTGAGAACCATGATTCC 600

QY 569 C-TTCCGTCCTCAAGCCCGAGGTTTCTCTGCCCCCCCAGCCCCCTGGACAA----TGCTGGCCT 622

Db 601 CTTTCCGTCCTCAAGCCCGAGGTTTCTCTGCCACCCAGGCCCTGGACAAGTGGTTGGGCATGT 660

QY 623 GTTCTCTACCTCACCGTGTCTATGGCTCACACCCCGCTCATGATCCAAAGCTTACGGAGTCG 682

Db 661 TCTTCGTAACTCACCGTGTCTATGGTCACCCCGCTCATGAATCCAAAGCTTACAGGAGTCG 720

QY 683 CTTAGAT-GAGAACACCATCCCTCCACTGTCCAGTCCATG---ATGCTCAGACAAAAATG 738

Db 721 GTTAGATGGANGCAACCATCCCTCCACTGTGAGGGCAGGGATTGGCTCGGACCAGAAATT 780

QY 739 TCCAAAGGCTTCACCGCCTTTTGGGAAGA 766

Db 781 GTCAGCGGTACACGGCTTTTGAGAAAAA 808

RESULT 13

BQ267358

LOCUS BQ267358 578 bp mRNA linear EST 07-MAY-2002

DEFINITION ik02g05.y1 Human insulinoma Homo sapiens cDNA clone IMAGE: 5779785 5', similar to TR:O14517 O14517 SMRP. [3] TR:Q9UNP5 ;, mRNA sequence.

ACCESSION BQ267358

VERSION BQ267358.1 GI:20492430

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this is a Washington University Pancreas EST project library. "

BASE COUNT 153 a 137 c 146 g 131 t

ORIGIN

Query Match 5.1%; Score 249; DB 13; Length 567;

Best Local Similarity 65.6%; Pred. No. 1.4e-60;

Matches 363; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 3802 GAGACAACACACCCACCGTGTCTTACGGCATCAACCTGACCATCCGCGCCACGAAGTGG 3861

Db 8 GAGAAAACCTCCCTCTCGTCCTAAAGAAAGTATCCTTCACGATCAAACTTAAAGAGAAGA 67

QY 3862 TGGGCATCGTGGGAAGGACGGGCTCTGGGAAGTCCCTTGGGCATGGCTCTTCCGCC 3921

Db 68 TTGGCATTGTGGCGGACAGGATCAGGGAAGTCTCGCTGGGATGGCCCTCTTCCGTC 127

QY 3922 TGGTGGAGCCCATGGCAGGCCGGATTCTCATTTGACGGCGTGGACATTTGCAGCATCGGCC 3981

Db 128 TGGTGGAGTTATCTGGAGGCTGCATCAAGATTGATGGAGTGAGATCAGTGATATTGGCC 187

QY 3982 TGGAGGACTTGGGTCCAAGCTCTCAGTGATCCCTCAAGATCCAGTCCAGTCTCTCAGGAA 4041

Db 188 TTGCCGACCTCCGAAGCAAACTCTCTATCATCTCCTCAAGAGCCGGTGTTCAGTGGCA 247

QY 4042 CCATCAGATTTCAACCTAGATCCCTTTGACCGTCACACTGACCAGCAGATCTGGGATGCCT 4101

Db 248 CTGTCAGATCAAAATTTGGACCCCTTCAACCCAGTACACTGAAGACCAGATTTGGGATGCC 307

QY 4102 TGGAGAGGACATTCCTTGACCAAGGCCATCTCAAAGTTCCCCAAAAAGCTGCATACAGATG 4161

Db 308 TGGAGAGGACACACATGAAAGAAATGTATTGCTCAGCTACCTCTGAAACTTGAATCTGAAG 367

QY 4162 TGGTGGAAAAACGTGGAAACTTCTCTGTGGGGGAGAGGCAGCTGCTGTGCATTGCCAGGG 4221

Db 368 TGATGGAGAAATGGGGATAACTTCTCAGTGGGGGAACGGCAGCTCTTGTGCATAGTAGAG 427

QY 4222 CTGTGCTTCGCAACTCCAAGATCATCCTTATCGATGAAGCCACAGCCCTCCATTGACATGG 4281

Db 428 CCCTGCTCCGCCACTGTAAGATTCTGATTTAGATGAAGCCACAGCTGCCATGGACACAG 487

QY 4282 AGACAGACACCCGTATCCAGCGCACAAATCCGTGAAGCCCTTCCAGGGGTGCACCGTGTCTCG 4341

Db 488 AGACAGACTTATTGATTCAAGAGACCCATCCGAGAAGCATTTGCAGACTGTACCATGCTGA 547

QY 4342 TCATTGCCCCACCG 4354

Db 548 CCATTGCCCATCG 560

RESULT 15

BQ687464

LOCUS

DEFINITION BQ687464 888 bp mRNA linear EST 15-JUL-2002

AGENCOURT_8058516 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6206756

5', mRNA sequence.

ACCESSION BQ687464

VERSION BQ687464.1 GI:21812780

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 888)

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LCM2361 row: 1 column: 21

High quality sequence stop: 707.

FEATURES

source

Location/Qualifiers

1..888

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6206756"

/clone_lib="NIH_MGC_110"

/tissue_type="ductal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: pancreas; Vector: pOFB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 178 a 263 c 260 g 186 t

ORIGIN

Query Match 4.7%; Score 230.4; DB 14; Length 888;

Best Local Similarity 57.5%; Pred. No. 5.4e-55;

Matches 414; Conservative 0; Mismatches 306; Indels 0; Gaps 0;

QY 3709 CGGAAGCTCCTTTACACATGGAAGGCACAAGTTGTCCCCAGGGGTGGCCACAGCATGGGG 3768

Db 31 CAGAGGGCCCTGGGTGGTGAAGGCAGCGCCCTCCCCGAAGTTGGCCCCCACGTGGGG 90

QY 3769 AAATCATATTTTCAGGATTATCACATGAAATACAGAGACAACACACCCACCCTGCTTCACG 3828

Db 91 AGGTGGAGTTCGGGAATTATTCTGTGCGCTACCGCGCGGCTAGACCTGGTGTGAGAG 150

QY 3829 GCATCAACCTGACCATCCCGGCCACGAAAGTGGTGGGCATCGTGGGAAGGACGGGCTCTG 3888

Db 151 ACCTGAGTCTGCATGTGCACGGTGGCGAGAAGTGGGATCGTGGCGCCGACTGGGGCTG 210

QY 3889 GGAAGTCCCTTGGGCATGGCTCTCTCCGCTTGGTGGAGCCCATGGCAGGCCGGATTTC 3948

Db 211 GCAAGTCTCCATGACCCCTTTGCCTGTTCGCGATCCTGGAGGGCGCAAGGGTGAATCC 270

QY 3949 TCATTGACGGCGTGGACATTTGCAGCATCGGCCCTGGAGGACTTGGGTGCCAAGCTCTCAG 4008

Db 271 GCATTGATGGCCTCAATGTGGCAGACATGGCCTCCATGACCTGGCTCTCAGCTGACCA 330

QY 4009 TGATCCCTCAAGATCCAGTGTGCTCTCAGGAACCATCAGATTCACACCTAGATCCCTTTG 4068

Db 331 TCATCCCGCAGGACCCCATCCTGTTCTCGGGGACCCTTCGCGCATGAACCTGGACCCCTTCG 390

QY 4069 ACCGTACACTGACCAGCAGATCTGGATGCCCTTGGAGAGGACATTCCTGACCAAGGCCA 4128

Db 391 GCAGCTACTCAGAGGAGGACATTTGGTGGGTTTGGAGCTGTCCCACCTGCACACGTTTG 450

QY 4129 TCTCAAAGTTCCCCAAAAGCTGCATACAGATGGTGGAAAACGGTGGAAACTTCTCTG 4188

Db 451 TGAGCTCCAGCCGGCAGGCCCTGGACTTCCAGTGCTCAGAGGGGGGAGAGATCTCAGCG 510

QY 4189 TGGGGGACAGGCAGCTGCTCTGCATTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCC 4248

Db 511 TGGGGCAGAGGCAGCTCGTGTGCCTGGCCCGAGCCCTGCTCCGCAAGAGCCGCTCTCTGG 570

QY 4249 TTATCGATGAAGCCACAGCCCTCCATTGACATGGAGACAGACACCCCTGATCCAGGCACAA 4308

Db 571 TTTTAGACGAGGCCACAGCTGCCATCGACCTGGAGACTGACAAACCTCATCCAGGCTACCA 630

QY 4309 TCCGTGAAGCCTTCCAGGGCTGCACCGTGTCTGTCATTGCCCCACCCTGTCAACCACTGTGC 4368

Db 631 TCCGCACCCAGTTTGATACCTGCACCTGTCTCTGACCATCGCACACCGGCTTAACACTATCA 690

QY 4369 TGAACCTGTGACCACATCCTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCGG 4428

Db 691 TGGACTACACCAGGGTCCTGGTCCTGGACAAAGGAGTAGCTGAATTTGATTCTCTCCAG 750

Search completed: July 24, 2003, 11:14:39
Job time : 6178 secs

THIS PAGE BLANK (USPTO)